Analyse probabiliste de protocoles de population

Yves Mocquard

UNIVERSITE BRETAGNE LOIRE / MATHSTIC

UNIVERSITÉ DE RENNES 1

THÈSE DE DOCTORAT DE

L'UNIVERSITE DE RENNES 1
COMUE **U**NIVERSITE **B**RETAGNE **L**OIRE

Ecole Doctorale N° 601
Mathématique et Sciences et Technologies de l'Information et de la Communication (MathSTIC)
Spécialité : Informatique

Par

Yves MOCQUARD

« Analyse probabiliste de protocoles de population »

Thèse présentée et soutenue à RENNES, le 13 décembre 2018
Unité de recherche : UMR 6074, IRISA, équipe Dionysos
Thèse N° :

Rapporteurs avant soutenance :

Johanne COHEN	Directrice de recherche CNRS au LRI - Université Paris Sud
Landy RABEHASAINA	Maître de conférences HDR au Laboratoire de Mathématiques de Besançon

Composition du jury :

Président :	**François CASTELLA**	Professeur à l'IRMAR - Université de Rennes 1
Examinateurs :	**Jérémie CHALOPIN**	Chargé de recherche CNRS au LIS - Marseille
	Johanne COHEN	Directrice de recherche CNRS au LRI - Université Paris Sud
	Landy RABEHASAINA	Maître de conférences HDR au Laboratoire de Mathématiques de Besançon
Dir. de thèse :	**Bruno SERICOLA**	Directeur de recherche INRIA - Rennes
Co-encadrante :	**Emmanuelle ANCEAUME**	Chargée de recherche CNRS à l'IRISA - Rennes

"J'ai rédigé dans le présent livre et je t'envoie les démonstrations de ces deux théorèmes. Mais te voyant, comme j'ai coutume de le dire, savant zélé, philosophe distingué et grand admirateur des recherches mathématiques, j'ai cru devoir y consigner également et te communiquer les particularités d'une certaine méthode dont, une fois maître, tu pourras prendre thème pour découvrir, par le moyen de la mécanique, certaines vérités mathématiques. Je me persuade d'ailleurs que cette méthode n'est pas moins utile pour la démonstration même des théorèmes. Souvent, en effet, j'ai découvert par la mécanique des propositions que j'ai ensuite démontrées par la géométrie — la méthode en question ne constituant pas une démonstration véritable. Car il est plus facile, une fois que par cette méthode on a acquis une certaine connaissance des questions, d'en imaginer ensuite la démonstration, que l'on recherchait sans aucune notion préalable." Archimède (287-212 av. J.-C.) Introduction à "La Méthode" adressée à Ératosthène (v. 276-194 av. J.-C.). Traduction T. Reinach, 1908.

Remerciements

Je tiens à remercier l'Irisa qui m'a accueilli pendant ces 3 années dans l'équipe de recherche Dionysos au sein de laquelle il est agréable de travailler dans une ambiance studieuse et chaleureuse.

Merci aux membres du jury, Johanne Cohen, Landy Rabehasaina, François Castella et Jérémie Chalopin pour l'intérêt qu'ils ont porté à mes travaux et pour leurs remarques pertinentes.

Je remercie particulièrement mes encadrants Emmanuelle Anceaume et Bruno Sericola. Ils m'ont encouragé et soutenu, me laissant une grande liberté tout en exigeant la rigueur indispensable à une démarche scientifique.

Merci à tous ceux qui m'ont accompagné à titres divers dans ce parcours : James Aspnes, Yann Busnel, Samantha Robert, Rumen Andonov, Romaric Ludinard, Corentin Hardy, Fabienne Cuyollaa, tous les membres de l'équipe Dionysos et mon entourage amical et familial.

Table des matières

Chapitre 1

Introduction

Les réseaux informatiques modernes sont de taille de plus en plus gigantesques et la distribution des calculs est un passage obligé. Cette distribution peut se contrôler de manière centralisée. Cela conduit à deux inconvénients : d'une part le système devient entièrement dépendant de la machine qui est au centre et contrôle la distribution, rendant le système extrêmement vulnérable à une attaque ou même tout simplement à une panne ; d'autre part les échanges se faisant essentiellement entre cette machine centrale et les autres, un goulot d'étranglement sur la bande passante existera obligatoirement au niveau de la machine centrale.

Avec un système distribué non centralisé, il est important d'obtenir des informations sur l'état global du système ce qui est loin d'être trivial en l'absence de centralisation.

Les protocoles de population offrent une solution en permettant l'émergence de certaines données globales du système uniquement au moyen d'interactions pair à pair aléatoires. Les données globales qui peuvent ainsi être collectées, sont la proportion, le nombre ou la majorité de noeuds dans un certain état, ou encore un horodatage du système où l'unité de temps est soit l'interaction par site, soit le temps d'une diffusion.

Les protocoles de population permettent l'envoi d'un message par diffusion, l'élection de leader ou l'élection de junte.

Avec les protocoles de population nous obtenons une modélisation qui peut s'appliquer sur des systèmes déjà définis. Par exemple, la diffusion de rumeur est un phénomène universel intéressant la sociologie, l'épidémiologie, le marketing et l'informatique distribuée, entre autres. Elle peut très simplement se modéliser à l'aide d'un protocole de population (voir chapitre 4).

Cette thèse, loin d'être exhaustive, présente quelques protocoles de population, en général assez simples, ainsi qu'une analyse probabiliste détaillée de chacun de ceux-ci.

Dans ce chapitre d'introduction, vont être présentés le déroulement chronologique de ces trois années de thèse, puis les détails de la méthodologie employée et enfin le contenu de chaque chapitre de ce manuscrit.

1.1 Déroulement de la thèse

1.1.1 Premier contact avec les protocoles de population

En 2015, dans le cadre de mon stage de Master 2 à l'Irisa sur les protocoles de population, j'ai lu un article de Aspnes et Ruppert [15] qui présente un protocole de majorité [15]. En effectuant des simulations, je me suis rendu compte que le protocole ne fonctionnait pas. Nous avons prouvé que la probabilité qu'il fonctionne était inférieure à $1/n!$, n étant la taille du système. Nous avons informé James Aspnes qui nous a indiqué de supprimer la dernière interaction. Effectivement, cette suppression a permis d'obtenir le protocole à quatre états déjà publié dans [33] et [51]. Ce fut le point de départ d'échanges fructueux sur nos travaux respectifs, particulièrement sur les protocoles basés sur la moyenne avec des nombres réels ou entiers. Chacun de notre coté, nous avons cherché et trouvé la démonstration du théorème 5.2.4. Ce théorème fournit un majorant de l'espérance d'une mesure basée sur la norme euclidienne du protocole de moyenne. Ceci nous a amené à écrire conjointement l'article [54].

1.1.2 Le protocole basé sur la moyenne

Dans le cadre de la thèse, les simulations montraient que le protocole de moyenne pouvait être amélioré en diminuant le nombre d'états nécessaires, sans que la performance en nombre d'interactions en pâtisse. En effet, au regard des simulations nous constatons qu'un protocole de moyenne avec des entiers converge vers un état où la différence maximale entre deux valeurs est inférieure ou égale à 2, en $O(n \log n)$ interactions. La recherche de la démonstration du théorème 5.2.12 a représenté beaucoup d'énergie, et ce théorème démontre exactement ce que nous avions constaté. Cela a été l'objet de l'article [55] fin 2016.

1.1.3 Le protocole de diffusion de rumeur

Au printemps 2016, nous avons travaillé sur la diffusion de rumeur. Des résultats nouveaux ont été obtenus sur la queue de distribution et sur la fonction de répartition du temps de convergence du protocole de diffusion de rumeur. Nous avions également des résultats intéressants sur le comportement de la fonction de répartition du temps de convergence autour de l'espérance quand on fait tendre n (la taille du système) vers l'infini. Cela a donné lieu à l'article [57].

1.1.4 Une mesure fructueuse

Dans [54], pour traiter le protocole de la moyenne avec des réels, nous avions utilisé comme mesure la norme euclidienne qui nous servait ensuite pour faire une approximation sur la norme infini. Assez naturellement, nous avons essayé d'utiliser la norme 4 et cela a donné de bons résultats même si les calculs se sont révélés être assez complexes. Cela a donné lieu à l'article [58] en novembre 2017.

1.1.5 Le protocole de diffusion de rumeur, suite

Nous avons poursuivi le travail sur la diffusion de rumeur en résolvant une conjecture que l'on avait émise et en étendant les résultats au temps continu. Nous avons publié ces résultats dans une revue internationale [60].

1.1.6 L'horloge globale

Dans tous nos travaux, nous avons donné des bornes très précises sur les temps de convergence, c'est-à-dire avec des constantes numériques explicites. L'idée sous-jacente à cette précision était de pouvoir coordonner plusieurs protocoles en estimant, à l'aide d'une mesure du temps, si la convergence avait eu lieu avec une probabilité élevée ou non. Une horloge en tant que protocole de population faisait défaut. Nous avions quelques idées pour concevoir un protocole d'horloge mais rien de probant. En 2017, Alistarh et al. [6] ont présenté un protocole de majorité qui utilisait une horloge dont le principe était extrêmement simple et bien étudié. Il se basait sur un article de Peres et al. [67]. Ce dernier article, qui avait une preuve assez complexe, mais malgré tout élégante, ne présentait pas de constantes. Nous nous sommes attelés à les exhiber en nous basant sur le même schéma de preuve mais en paramétrant certaines constantes de la preuve pour pouvoir optimiser. Ceci a donné lieu à la publication [56].

1.1.7 La détection de convergence

Nous avions tous les ingrédients pour écrire ce dernier article qui utilise le protocole basé sur la moyenne avec les entiers, le protocole de diffusion de rumeur et l'horloge globale. Pour chacun de ces protocoles, nous avons approfondi l'existant. Cela nous a permis de construire un mécanisme de détection de convergence. Ceci a donné lieu à la publication [59] en novembre 2018.

1.2 Méthodologie

À de nombreuses reprises, la méthodologie suivie pour le travail de recherche, fut similaire à celle décrite par Archimède dans la citation en exergue, en remplaçant "mécanique" par "simulation" et "géométrie" par "mathématiques".

1.2.1 Les simulations

Il s'agit de la simulation de protocoles de population, au moyen d'un programme informatique utilisant un générateur pseudo-aléatoire. Les protocoles que nous étudions fonctionnent avec des interactions aléatoires uniformément distribuées, chaque interaction entre agents consiste en un petit algorithme très simple. Cela se prête donc particulièrement bien à une simulation de type Monte-Carlo.

Les simulations répondent à plusieurs objectifs.

Dans le travail en amont, elles permettent de voir comment se comporte un protocole, afin d'orienter les recherches.

Les simulations permettent d'illustrer la pertinence de ce qui a été prouvé théoriquement.

Enfin, à l'aide de simulations, nous pouvons proposer des conjectures assez solides sur les bornes des temps de convergence.

Pour tout ce travail, nous nous sommes appuyés sur un logiciel que nous avons développé en Java. Chaque protocole est une classe implémentant une interface générique. Une autre classe se charge de simuler les interactions aléatoires. Il y a plusieurs modes, le mode simple où il s'agit de voir grossièrement et rapidement le comportement d'un protocole, et un mode élaboré où de nombreuses simulations sont effectuées en parallèle grâce à l'utilisation du multi-threading. Le code source de ce logiciel sera mis en ligne prochainement.

1.2.2 Les démonstrations

Les démonstrations sont le cœur de ce travail. Pour plus de clarté, les démonstrations un peu longues ont été placées en annexe afin de garder une meilleure lisibilité au reste du chapitre. Deux théorèmes et leur démonstration méritent d'être mis en avant, sans rien enlever au mérite des autres théorèmes.

Le théorème 5.2.12 avec sa démonstration en deux temps (5.2.9 et 5.2.11), est celui qui a été le plus laborieux. Les deux parties de ce théorème suivent de façon très différente des cheminements originaux qui se révèlent être efficaces.

Le théorème 4.2.4 donne un résultat précis sur la diffusion de rumeur qui est un processus assez universel. Ce résultat est assez surprenant du fait d'une formulation différente selon que le nombre nœuds informés à l'origine est ou non égal à 1. Le théorème 4.2.5 est une application directe de ce théorème. Les figures 4.5 et 4.6 en section 4.2.5 montrent à quel point les bornes théoriques issues de 4.2.4 sont proches des résultats expérimentaux.

1.3 Organisation du manuscrit

Le chapitre 2 présente le modèle des protocoles de population et définit tous les concepts utilisés par la suite. Les protocoles de population sont présentés de manière informelle avec une définition des différents termes utilisés. Puis, une présentation précise et formelle du modèle est donnée. La notation utilisée par la suite est explicitée. Les concepts de synchrone/asynchrone, le modèle de fautes ainsi que le temps discret/continu sont abordés.

Le chapitre 3 définit les problèmes que nous serons à même de résoudre dans les chapitres suivants. Pour chacun des problèmes, il y a 3 trois sous-sections : la définition du problème, l'état de l'art, et une présentation des résultats obtenus.

Le chapitre 4 étudie le problème de la diffusion de rumeur. Les deux sections principales traitent du temps discret pour l'une et du temps continu pour l'autre. Les résultats principaux sont l'expression explicite de la loi du temps de convergence, l'expression d'une borne qui est aussi un équivalent de cette loi, et enfin, une analyse fine du comportement asymptotique de ce temps de convergence autour de son espérance quand la taille du système n tend vers l'infini. Tout ceci est montré aussi bien en temps discret qu'en temps continu.

Le chapitre 5 traite du protocole de moyenne. La section 5.1 décrit des protocoles de moyenne avec des réels. L'originalité de ce travail est l'utilisation de la norme 4 qui permet d'avoir des bornes plus pertinentes. Les problèmes abordés sont la proportion, le comptage et la majorité. La section 5.2 aborde les protocoles de moyenne avec des entiers, principalement au travers du problème de la proportion. Le résultat principal est le théorème 5.2.12. Dans cette section nous traitons également le problème du comptage et le caractère optimal de la solution proposée aussi bien pour le comptage que pour la proportion.

Le chapitre 6 traite du protocole d'horloge globale. Il reprend le même schéma de démonstration que Peres et al. et Alistarh et al. [67, 10]. Nous allons plus loin en explicitant les bornes, ce qui n'avait jamais été fait auparavant.

Le chapitre 7 traite d'un mécanisme de détection de convergence qui utilise les résultats de l'horloge globale ainsi que ceux de la diffusion de rumeur. Un exemple est donné avec le protocole de calcul de proportion.

Le chapitre 8.1 est la suite du chapitre 3. Il définit des problèmes qui, tout en ayant leur importance dans le cadre des protocoles de population, n'ont pas été traités spécifiquement dans cette thèse.

Le chapitre 8 conclut en mettant en évidence les perspectives.

Après la bibliographie, plusieurs annexes sont dédiées aux démonstrations les plus longues.

Chapitre 2

Le modèle des protocoles de population

2.1 Les protocoles de population

Nous avons un ensemble de n **agents** que nous appelons aussi **nœuds**. Ces agents sont les sommets d'un **graphe d'interactions** G que nous supposons connexe. L'ensemble d'agents se nomme indifféremment **système** ou **réseau**. La **taille du système** est n.

Les agents interagissent séquentiellement deux à deux, l'interaction s'effectue entre deux voisins du graphe d'interactions G. À chaque instant, deux agents voisins sont choisis aléatoirement pour interagir. Les agents sont les sommets d'un graphe orienté complet G' dont le poids des arêtes représente les probabilités d'interaction. Pour les agents non voisins par G, le poids de l'arête les rejoignant par G' est nul.

Chaque agent possède trois types d'états.

Premièrement, l'**état d'entrée** est l'état de l'agent au démarrage du système.

Deuxièmement, l'**état de travail** est l'état de l'agent au moment de chaque interaction. C'est l'état le plus important, quand on parle d'**état** sans plus de précision, il s'agit de l'état de travail. Cet état est initialisé avec l'image par la **fonction d'entrée** de l'état d'entrée. L'état de travail est ensuite mis à jour lors de chaque interaction à laquelle l'agent participe. Cette mise à jour se fait en suivant une **fonction de transition**. La fonction de transition a pour variable d'entrée un couple d'états de travail et comme sortie un autre couple d'états de travail. Le premier couple (l'antécédent de la fonction) représente l'état des agents avant l'interaction. Le deuxième couple (l'image de la fonction) représente l'état des agents après l'interaction. Notons que lors d'une interaction, l'ordre des agents a une importance : le premier agent agit en tant qu'initiateur, le second en tant que répondeur [11]. La correspondance entre tous les agents et leur état respectif à un instant donné est appelé **configuration**.

Troisièmement, l'**état de sortie** est l'état rendu à l'utilisateur lorsqu'il interroge un agent. Cet état de sortie est l'image par la **fonction de sortie** de l'état de travail.

Le plus souvent, le but d'un protocole de population est d'obtenir des informations au moyen de l'interrogation d'un agent par la fonction de sortie sur l'état global du système à l'initialisation, c'est-à-dire sur la distribution des états d'entrée au dé-

marrage du système. Quand toutes les sorties donnent les informations attendues, on dit que le système est dans un état de **convergence**.

Pour d'autres protocoles, c'est une démarche inverse. À partir d'une configuration d'entrée bien définie, on cherche à atteindre une configuration particulière des états de sortie comme, par exemple, dans le cas d'un protocole d'élection de leader (voir section 8.1.1) où la convergence est atteinte quand il n'y a plus qu'un et un seul leader dans le système.

Le **temps de convergence** est le temps qu'il faut à partir de l'état initial pour atteindre la convergence. C'est une variable aléatoire.

2.2 Le modèle formel et conventions d'écriture

Voici le modèle formel tel que décrit par Angluin et al. dans [14], à partir duquel chaque protocole sera décrit. Dans la suite de cette section, nous décrivons les conventions d'écritures générales qui seront précisées dans chacun des chapitres traitant de protocole.

Définition 2.2.1 *Un* **protocole de population** *sur un graphe d'interaction complet se caractérise par un sextuplet* $(\Sigma, Q, \Xi, \iota, f, \omega)$ *comprenant :*

- *un* **ensemble d'entrée** *:* Σ *est l'ensemble fini des états d'entrée. C'est l'ensemble des états possibles des agents à l'origine.*
- *un* **ensemble d'états** *:* Q *est l'ensemble fini des états de travail de chaque agent. Ce sont ces états qui entrent en jeu lors d'une interaction entre deux agents.*
- *un* **ensemble de sortie** *:* Ξ *est l'ensemble fini des états de sortie. Quand on interroge un agent, le résultat est un élément de cet ensemble.*
- *une* **fonction d'entrée** $\iota : \Sigma \to Q$*, est la fonction qui permet d'initialiser l'état de travail à partir de l'état d'entrée.*
- *une* **fonction de transition** $f : Q \times Q \to Q \times Q$*, fonction de transition, elle définit la mise à jour des états de chaque agent lors d'une interaction.*
- *une* **fonction de sortie** $\omega : Q \to \Xi$*, fonction qui donne le résultat de l'interrogation d'un agent. C'est une fonction de l'état de travail.*

Chaque sommet du **graphe d'interaction** *est un agent et chaque arête a pour poids la probabilité d'interaction entre deux agents.*

Initialement, chaque agent possède un symbole de Σ, la fonction d'entrée permet, en fonction de ce symbole, d'initialiser l'état de chaque agent avec un élément de Q. Puis, lors des interactions entre les agents, cet état est mis à jour selon la fonction de transition f. A tout instant, on peut interroger un agent qui répondra par un élément de Ξ, cet élément est l'image par la fonction ω de l'état de cet agent.

Les interactions entre agents sont orchestrées par un ordonnanceur aléatoire : à chaque instant discret, deux agents sont choisis au hasard pour interagir.

La notion de temps dans les protocoles de population désigne le nombre d'interactions global, tandis que le temps parallèle désigne la moyenne du nombre d'interactions initiées par chaque agent [15]. Le temps parallèle est donc le temps global divisé par la taille n du système.

Les agents sont anonymes, ils ne gèrent ni n'utilisent d'identifiant. Cependant, pour faciliter la présentation, les agents sont numérotés de 1 à n. Pour conserver la notion d'anonymat, il va de soi qu'aucun agent ne connaît son numéro.

Les agents n'ont aucune connaissance directe du temps. Cependant, ils sont représentés par une composante d'un vecteur indicé par le temps t qui représente le nombre total d'interactions ayant eu lieu dans le système depuis l'origine.

Les états de chaque agent à l'instant discret $t \geq 0$ sont représentés par le vecteur $C_t \in Q^n$, c'est-à-dire que $C_t = \left(C_t^{(1)}, \ldots, C_t^{(n)}\right)$. Pour $i \in [\![1, n]\!]$ et $t \in [\![0, \infty[\![$, $C_t^{(i)} \in Q$ est l'état de l'agent i à l'instant t.

$C = \{C_t,\ t \geq 0\}$ est donc un processus stochastique dont les états sont éléments de Q^n.

A chaque instant discret t, deux agents distincts i et j sont successivement choisis parmi les nombres de 1 à n avec probabilité $p_{i,j}(t)$.

Le choix des deux agents interagissant est indépendant de leur état ainsi que de celui des autres agents. Autrement dit, nous supposons que la variable aléatoire décidant du choix des agents interagissant est indépendante de l'état du processus stochastique avant le choix.

Si i et j sont choisis pour interagir, alors le passage de la configuration C_t à C_{t+1} se fait ainsi

$$\left(C_{t+1}^{(i)}, C_{t+1}^{(j)}\right) = f\left(C_t^{(i)}, C_t^{(j)}\right) \text{ et } C_{t+1}^{(r)} = C_t^{(r)} \text{ pour } r \neq i, j \text{ et } r \in [\![1, n]\!].$$

Si la probabilité que les deux nœuds i et j, avec $i, j \in [\![1, n]\!]$ interagissent est indépendante du temps, nous notons cette probabilité $p_{i,j}$.

Dans notre définition des protocoles de population en début de cette section, nous avons parlé du graphe d'interaction en précisant qu'il était complet. Il est possible d'étendre cette définition aux graphes non complets. La matrice dont les coefficients sont $p_{i,j}$, est une manière élégante et précise de définir ce graphe, les arêtes n'existant pas étant représentées par une probabilité nulle d'interaction. Cette matrice représente le graphe d'interaction.

Dans tous les protocoles que nous étudions, comme c'est le cas dans la majorité des travaux sur les protocoles de population, nous supposons que le ordonnanceur suit une loi uniforme sur un graphe complet, c'est-à-dire que pour tout $i, j \in [\![1, n]\!]$

$$p_{i,j} = \frac{1_{i \neq j}}{n(n-1)}. \tag{2.1}$$

Le but d'un protocole de population est d'atteindre un état de convergence, cet état de convergence signifie que l'ensemble des résultats des fonctions de sortie de chaque agent vérifie certaines propriétés qui dépendent du problème que cherche à résoudre le protocole. Nous en donnerons quelques exemples plus loin.

Rappelons que la mesure du temps, dans un protocole de population, est le nombre d'interactions et que le temps parallèle est le temps divisé par n.

2.3 Performance d'un protocole de population

Un problème étant posé, il peut y avoir plusieurs protocoles permettant de le résoudre. Dans cette section, nous examinons les deux critères fondamentaux qui mesurent la performance de ces protocoles. Ces deux critères, par analogie avec les problèmes algorithmiques classiques, sont nommés le temps et l'espace.

2.3.1 Le temps de convergence

Le temps de convergence est, comme nous l'avons déjà dit, une variable aléatoire. Pour un protocole, ce temps correspond à la durée nécessaire pour atteindre la convergence qui est son objectif. Comme en algorithmique classique, plus ce temps est réduit plus le protocole est performant.

Deux indicateurs sur le temps de convergence nous permettent de cerner les performances temporelles d'un protocole de population.

- Le **temps de convergence moyen** ou **espérance du temps de convergence**

- Le **temps de convergence avec probabilité élevée** est le temps après lequel la probabilité de convergence est supérieure à $1-1/n$, n étant le nombre d'agents.

Le temps de convergence avec probabilité élevée possède deux autres formulations plus précises. Pour l'une il s'agit, $\sigma > 0$ étant donné, du temps après lequel la probabilité de convergence est supérieure à $1 - 1/n^\sigma$. Pour l'autre, $\delta \in]0;1[$ étant donné, il s'agit du temps après lequel la probabilité de convergence est supérieure à $1-\delta$. Il est clair que ces deux dernières formes sont équivalentes et elles impliquent la première (en prenant $\sigma = 1$ ou $\delta = 1/n$). Nous préférons utiliser la formulation avec δ dans la mesure où cela donne une information plus fine sur le temps de convergence.

2.3.2 Le nombre d'états

Le plus souvent, l'ensemble des états d'entrée Σ et l'ensemble des états de sortie Ξ sont des données du problème, le reste du protocole (Q, ι, f et ω) étant la solution.

Le nombre d'états $|Q|$ d'un protocole de population étant fini, plus sa taille est réduite, plus nous considérons que le protocole est performant.

Il y a souvent un compromis avec le temps de convergence. Il existe des protocoles qui privilégient un temps de convergence réduit aux dépens du nombre d'états, pour d'autres c'est l'inverse et, enfin, il existe certains protocoles où un équilibre est cherché entre ces deux critères de performance.

2.4 Précisions sur le modèle

Cette section a pour but d'introduire certaines notions importantes de l'informatique distribuée qui seront évoquées par la suite, principalement dans les sections consacrées à l'état de l'art.

2.4.1 Push, pull et push/pull

Les termes push (pousser), pull (tirer) et push/pull correspondent à des protocoles n'utilisant qu'un certain type d'interaction. Dans un protocole où les interactions sont de type push, l'agent à l'initiative d'une interaction reste invariant. Dans un protocole où les interactions sont de type pull, c'est le répondeur qui reste invariant. Dans un protocole où les interactions sont de type push/pull, il n'y a aucune obligation de rester invariant en fonction du rôle dans l'interaction. Les protocoles push et pull, parfois qualifiés d'unidirectionnels [12], peuvent être assimilés à des protocoles où les interactions se font respectivement par envoi et réception de messages. Certains protocoles s'y prêtent mieux que d'autres, par exemple le protocole diffusion de rumeur (voir chapitre 4), de nombreux protocoles d'élection de leader (voir section 8.1.1) ou encore le protocole d'élection de junte (voir section 8.1.2). Par contre, les protocoles basés sur la moyenne (voir chapitre 5) doivent impérativement s'exécuter dans un mode push/pull.

2.4.2 Synchrone/asynchrone

Le modèle synchrone ou asynchrone est lié à la maîtrise du temps.

Si les interactions ont lieu à des instants aléatoires et/ou avec des durées non maîtrisées, la seule contrainte étant qu'un agent ne peut interagir qu'avec un seul agent à la fois, alors nous sommes dans un mode asynchrone. Si les noeuds ont une horloge commune et que les interactions s'exécutent par rondes regroupant un certain nombre d'agents, nous sommes en mode synchrone.

Les protocoles de population ont été conçus pour modéliser des interactions distribuées entre agents. Une interaction est considérée comme ponctuelle dans le temps, c'est-à-dire que sa durée est nulle. De ce fait, deux interactions ne peuvent avoir lieu au même moment. Elles ont lieu de manière séquentielle. Cela correspond à un modèle asynchrone. Nous pourrions modéliser un modèle synchrone avec un protocoles de population. Cela consisterait à ce que l'ordonnanceur regroupe les interactions par rondes. Une ronde est un ensemble de couples d'agents, parmi lesquels on ne trouve pas plusieurs fois le même agent. A notre connaissance, le modèle synchrone n'a jamais été étudié dans le cadre des protocoles de populations. Il l'a été davantage par ailleurs [9, 36, 69, 46, 40].

2.4.3 Modèle de fautes

Le modèle de fautes est une notion importante du calcul distribué. La faute la plus courante est la non-réponse. Dans le cadre de protocole de population, cela se traduit par un noeud qui n'interagit plus ou qui interagit moins. Il est possible

de modéliser ce dernier comportement au niveau de l'ordonnanceur. Les erreurs de comportement qui sont plus graves et plus difficiles à contrer, se rapportent au comportement byzantin, comportement qui ne correspond pas à ce qui est prévu. Le terme byzantin vient d'un article célèbre de Lamport et al. [49] paru en 1982, il faisait référence à un siège où des généraux byzantins devaient prendre une décision alors que certains d'entre eux, de manière cachée, étaient liés à l'ennemi. Des protocoles admettent une certaine proportion d'agents au comportement byzantin. Dans le cadre des protocoles de population, un nœud byzantin ne respecte pas la fonction de transition et/ou induit en erreur le nœud avec lequel il interagit, en mentant sur son état. C'est une notion peu évoquée dans le cadre des protocoles de population, mais qui existe [13]. Aucun des protocoles présentés dans cette thèse n'est prévu pour fonctionner avec des nœuds byzantins. Nous introduisons cette notion car, dans l'état de l'art, nous faisons référence à des protocoles prenant en compte ce paramètre. Il existe d'autres types de fautes qui peuvent être considérés comme des cas particuliers d'erreurs byzantines. Les protocoles que nous présentons, sont prévus pour fonctionner dans un environnement exempt de toutes fautes.

2.4.4 Temps discret/continu

Comme nous l'avons vu un protocole de population fonctionne avec un temps discret, la mesure du temps étant le nombre d'interactions. Si nous associons à chaque agent un processus de Poisson et qu'à chaque saut, le nœud associé à ce processus choisisse un autre nœud pour interagir, alors à partir d'un protocole de population en temps discret, nous avons construit un protocole en temps continu. Tous nos protocoles sont en temps discret, le temps continu n'est étudié que pour la diffusion de rumeur.

Chapitre 3

Problèmes étudiés

Dans ce chapitre nous étudierons les problèmes solubles par des protocoles de population ayant un lien avec les différents articles publiés. D'autres problèmes seront évoqués chapitre 8.1.

Nous tenons à attirer l'attention sur le fait que dans la mesure où, dans ce chapitre, il s'agit de définir des problèmes, les protocoles ne sont vus que de l'extérieur. De chaque protocole $(\Sigma, Q, \Xi, \iota, f, \omega)$, nous ne décrivons précisément que l'ensemble d'entrée Σ et l'ensemble de sortie Ξ. Nous évoquerons la fonction de sortie sans la définir, cette fonction de sortie est appliquée à l'état d'un nœud. Nous rappelons que les états de tous les nœuds du système, c'est-à-dire la configuration, sont définis par le vecteur $C_t = (C_t^{(1)}, C_t^{(2)}, \ldots, C_t^{(n)}) \in Q^n$.

3.1 La diffusion de rumeur

3.1.1 Définition

Ce sujet est traité au chapitre 4. Contrairement aux autres problèmes, le protocole de diffusion de rumeur est défini dans sa version de base (nous verrons dans l'état de l'art qu'il en existe des variantes). À l'origine, un nombre non nul de nœuds sont initialisés à 1 (les nœuds "connaissant la rumeur" ou "diffuseurs") et les autres sont initialisés à 0 (les nœuds "ignorant la rumeur" ou "ignorants"). Lorsqu'un 0 interagit avec un 1 ou lorsqu'un 1 interagit avec un 0, les deux nœuds prennent la valeur 1, les autres interactions sont invariantes. La convergence est atteinte lorsque tous les nœuds connaissent la rumeur, c'est-à-dire quand tous les nœuds ont la valeur 1. Dans la mesure où le choix des interactions suit une loi uniforme, il est évident que le temps de convergence ne dépendra que du nombre de nœuds initialisés à 1 à l'origine. Dans le chapitre 4 dédié à la diffusion de rumeur, le nombre de nœuds ayant la valeur 1 à l'instant t est noté par Y_t, c'est-à-dire $Y_t = \sum_{i=1}^{n} C_t^{(i)}$.

Définition 3.1.1 *Soit $\delta \in]0, 1[$ et $i \in [\![1, n]\!]$. Un protocole de diffusion de rumeur commençant avec i nœuds connaissant la rumeur converge à l'instant $\tau \geq 0$ avec la probabilité $1 - \delta$, lorsque, pour tout $t \geq \tau$,*

$$\mathbb{P}\left\{Y_t = n \mid Y_0 = i\right\} \geq 1 - \delta$$

3.1.2 État de l'art

La diffusion de rumeur de manière aléatoire est un mécanisme important souvent utilisé en informatique, qui permet une diffusion d'informations dans de vastes et complexes réseaux au moyen d'interactions pair à pair. Ce mécanisme initialement proposé par Deemers et al [31] pour la mise à jour d'une base de données répliquée sur différents sites, a été adopté par différentes applications, de la découverte de ressources [43] aux applications distribuées complexes [24], en passant par l'agrégation de données [47].

La diffusion de rumeur est aussi un phénomène réel que l'on peut chercher à modéliser. Ce modèle peut avoir son utilité dans toutes sortes de domaines. Par exemple l'épidémiologie, avec l'étude de la diffusion de maladies infectieuses (Kermack et al. 1927 [48]), a été historiquement la première science à s'intéresser au phénomène. Plus récemment la diffusion de rumeur a été utilisée pour la modélisation de la propagation de virus dans un réseau informatique [21].

De nombreux travaux se sont concentrés sur la modélisation et l'étude de mécanismes de diffusion de rumeur aléatoire. Ces modèles peuvent être différents du modèle des protocoles de population, cependant la question principale reste toujours la même, celle du temps de convergence, c'est-à-dire le temps nécessaire pour que tous les nœuds du système soient informés de la rumeur.

Plusieurs modèles ont été pris en considération pour répondre à cette question. Le plus étudié est le modèle synchrone "push/pull" (pousser/tirer), aussi nommé appel téléphonique aléatoire synchrone. Ce modèle part de l'hypothèse que tous les sommets du graphe agissent de manière synchrone, ce qui permet de diviser le temps en tours synchronisés. Durant chacun de ces tours synchronisés, chaque sommet appelle un autre sommet choisi aléatoirement parmi ses voisins, s'il connaît la rumeur mais pas son voisin, il apprend **à** ce voisin la rumeur (opération "push"), s'il ne la connaît pas déjà mais que son voisin la connaît, il apprend **de** ce voisin la rumeur (opération "pull"). Dans le modèle synchrone, le temps de diffusion est le nombre de tours nécessaires pour que tous les sommets du graphe connaissent la rumeur.

Ce problème ("Telephone call problem") a été exposé une des toutes premières fois par Frieze et Grimmet [36], dans une version uniquement "push" et sur un graphe complet. Ils ont prouvé que, dans ces conditions, le ratio entre le temps de diffusion et $\log_2(n)$ converge en probabilité vers $1 + \ln(2)$, quand le nombre n de nœuds du graphe tend vers l'infini.

D'autre résultats ont aussi été établis (par exemple Pittel [69] en 1987 et Karp [46] en 2000), les résultats les plus récents viennent de l'observation selon laquelle le temps de diffusion de la rumeur est dépendante de la conductance du graphe associé au réseau, voir [40]. Des investigations ont aussi été conduites dans différentes topologies du réseau [25, 30, 35, 63] et dans la présence où non de nœuds défaillants [34].

Dans un réseau distribué, et en particulier dans un réseau distribué large échelle, l'hypothèse selon laquelle les nœuds agissent de manière synchronisée n'est pas réaliste. Récemment, un certain nombre d'auteurs se sont affranchis de cette hypothèse en considérant un modèle asynchrone. Dans le cas discret Acan et al. [1] ont étu-

dié le temps de diffusion pour de nombreuses topologies de graphe, en comparant le cas synchrone et le cas asynchrone. Il faut noter que le temps asynchrone pour Acan et al. est du temps continu par l'utilisation d'un processus de poisson associé à chaque nœud. Ils l'appellent horloge aléatoire exponentielle car dans tout processus de Poisson le temps entre deux sauts suit une loi exponentielle. Ils définissent le temps de diffusion garanti comme étant le temps minimum où, avec une probabilité d'au moins $1 - 1/n$, chaque nœud connaît la rumeur et aussi le temps de diffusion moyen qui est tout simplement l'espérance du temps de diffusion (on reconnaîtra la différence entre temps de convergence avec probabilité élevée et l'espérance du temps de convergence, faite dans la section 2.3.1). Ils montrent que le temps de diffusion moyen et celui avec probabilité élevée est $\Omega(n \log n)$, où n est le nombre de nœuds dans le réseau.

Angluin et al. [12] analysent le temps de propagation d'une rumeur en considérant seulement l'opération push (qu'ils appellent opération épidémique unidirectionnelle), et montrent qu'une rumeur injectée à un nœud nécessite $O(n \log n)$ interactions pour se diffuser à tous les nœuds du réseau, avec probabilité élevée. Ce résultat est intéressant, néanmoins les constantes découlant de la complexité ne sont pas déterminées. Dans le cas du temps continu, Ganesh [37] considère la propagation d'une rumeur quand il y a n processus de Poisson unitaires indépendants, chacun associé à un nœud. Au moment où il y a un saut du processus de Poisson associé au nœud i, ce nœud devient actif et choisit un autre nœud j uniformément au hasard parmi les autres pour communiquer. Ganesh [37] analyse l'espérance et la variance du temps de propagation de la rumeur sur les graphes généraux et Panagiotou et Speidel [64] proposent une étude approfondie de la diffusion de rumeur sur des graphes d'Erdös-Rényi (sur un graphe d'Erdös-Rényi la présence ou non de chaque arête est décidé par une variable aléatoire suivant une loi de Bernouilli de paramètre $p \in]0; 1[$) avec $p > c \ln(n)/n$ pour $c > 1$.

Dans [29], les auteurs proposent un modèle différent dans lequel, en plus des diffuseurs et des ignorants, est introduite la notion d'étouffeur. Un étouffeur connaît la rumeur mais ne la propage pas. Un étouffeur résulte de l'interaction entre deux diffuseurs, ou entre un diffuseur et un étouffeur. Ces auteurs ont conjecturé que le nombre d'étouffeurs suit de manière asymptotique une loi normale, l'espérance et la variance étant linéaires vis-à-vis de la taille du système. Ce modèle a été généralisé par Lebensztayn et al. dans un article [50] où les auteurs supposent en outre que chaque diffuseur cesse de propager la rumeur juste après avoir été impliqué dans un nombre aléatoire d'interactions avec des étouffeurs. Sous une configuration initiale générale, ils établissent le comportement asymptotique de la proportion finale d'ignorants en faisant tendre la taille du système vers l'infini. Dans [26], les auteurs proposent un modèle dans lequel les diffuseurs ont un capital d'émission aléatoire qui diminue à chaque émission. Ils étudient la proportion d'ignorants qui reçoivent l'information avant que le capital d'émission de tous les diffuseurs soit épuisé, ainsi que le temps d'épuisement. Ce travail a été étendu aux graphes d'Erdös-Rényi dans [27].

3.1.3 Nos résultats

Notre travail s'est concentré sur le modèle classique. En découlent naturellement des expressions explicites de l'espérance et la variance du temps de propagation T_n. Sachant que pour $t \geq 0$, $Y_t = \sum_{i=1}^{n} C_t^{(i)}$, T_n est défini par

$$T_n = \inf \{t \geq 0 \mid Y_t = n\}.$$

A partir des récurrences venant de l'analyse probabiliste de la diffusion, nous avons obtenu la loi de T_n. Nous avons obtenu un équivalent de la queue de distribution de cette loi qui se révèle aussi être une borne. Cet équivalent et cette borne possèdent un grand intérêt pratique. C'est d'ailleurs à partir de celle-ci que nous avons trouvé une expression du temps de convergence à partir de deux diffuseurs. Cette expression se révèle être remarquablement proche des résultats de simulations (voir section 4.2.5).

3.2 Le comptage

3.2.1 Définition

Ce sujet est traité dans le chapitre 5. Le comptage consiste, à partir d'un ensemble d'agents qui possèdent chacun une donnée d'entrée venant de l'ensemble $\Sigma = \{A, B\}$, à déterminer le nombre d'agents ayant démarré avec le symbole A. En termes plus formels si l'on note n_A (respectivement n_B), le nombre d'agents possédant le symbole A (respectivement B), le but d'un protocole de comptage est que chaque agent puisse déterminer, à l'aide de sa fonction de sortie ω_A, la valeur exacte n_A.

Définition 3.2.1 *Soit $\delta, \varepsilon \in]0, 1[$. Un protocole de comptage converge à l'instant $\tau \geq 0$ avec la probabilité $1 - \delta$, lorsque, pour tout $t \geq \tau$,*

$$\mathbb{P}\left\{\omega_A\left(C_t^{(i)}\right) = n_A \text{ pour tout } i \in [\![1, n]\!]\right\} \geq 1 - \delta.$$

3.2.2 État de l'art

Le problème du comptage n'a, à notre connaissance, jamais été abordé avant notre article [54] paru en 2015 puis [55] en 2016. Cependant la méthode que nous utilisons, c'est-à-dire la moyenne, a été beaucoup étudiée, principalement avec des réels, dans d'autres cadres que les protocoles de population. Dans [45] le modèle est synchrone et si les résultats peuvent être parfois similaires, ils n'ont pas les mêmes bases théoriques. Dongsheng et al.[65], dans le cadre du gossip asynchrone et en utilisant une mesure similaire à la norme euclidienne, démontrent dans leur lemme I un résultat quasi identique au corollaire 3 de notre article [54]. En 2012, Sauerwald et Sun [71] ont étudié le lien entre le protocole basé sur la moyenne avec des nombres réels et celui avec des entiers. Leur méthode consiste à mémoriser dans une matrice la différence entre les deux résultats (entiers et réels) sur des interactions identiques.

3.2.3 Nos résultats

En 2015, nous avons défini un protocole qui nécessitait un nombre d'états égal à $O(n^{3/2}/\sqrt{\delta})$ et un temps parallèle de $O(\log(n/\delta))$ pour converger [54]. En 2016, avec un protocole similaire et surtout un approfondissement des preuves, le nombre d'états nécessaires est devenu $O(n)$, le temps parallèle nécessaire pour converger est resté à $O(\log(n/\delta))$. De plus, nous avons prouvé que ces deux valeurs sont optimales [55].

Dans le cadre des protocoles de moyenne avec des réels, les résultats de notre publication [58] améliorent ceux de [54] par l'utilisation de la norme 4 plutôt que la norme euclidienne. Ils sont présentés dans la section 5.1.

3.3 Le problème de la proportion

3.3.1 Définition

Ce sujet est traité dans le chapitre 5. La proportion consiste, à partir d'un ensemble d'agents qui possèdent chacun une donnée d'entrée venant de l'ensemble $\Sigma = \{A, B\}$, à pouvoir déterminer la proportion d'agents ayant démarré avec le symbole A. En termes plus formels, on note n_A (respectivement n_B), le nombre d'agents possédant le symbole A (respectivement B) et γ_A (respectivement γ_B) la proportion d'agent possédant le symbole A (respectivement B), c'est à dire $\gamma_A = n_A/n$ (respectivement $\gamma_B = n_B/n$). Le but d'un protocole résolvant le problème de la proportion est que chaque agent puisse rendre, à l'aide de sa fonction de sortie ω_A, une valeur approchée de γ_A à ε près.

Définition 3.3.1 *Soit $\delta, \varepsilon \in]0,1[$. Un protocole de proportion avec la précision ε converge à l'instant $\tau \geq 0$ avec la probabilité $1-\delta$, lorsque, pour tout $t \geq \tau$,*

$$\mathbb{P}\left\{\left|\omega_A\left(C_t^{(i)}\right) - \frac{n_A}{n}\right| \leq \varepsilon \text{ pour tout } i \in [\![1,n]\!]\right\} \geq 1-\delta$$

3.3.2 État de l'art

Le problème de la proportion n'a, à notre connaissance, jamais été abordé avant notre article [55] paru en 2016 puis [58] en 2017. Il s'agit du même type de protocole que celui du comptage vu précédemment.

Il ne faut confondre la proportion avec un autre type de problème qui sera décrit dans la section 8.1.5, la proportion en tant que résultat.

3.3.3 Nos résultats

Nous avons démontré qu'avec un nombre d'états égal à $\lceil 3/(2\varepsilon) \rceil$, le protocole converge en un temps parallèle de $O(\log[n/(\varepsilon\delta)])$, ceci a été publié en 2016 [55], puis en 2018 [59]. Il faut noter que notre protocole ne dépend que de ε. Une fois ε fixé le même protocole peut fonctionner avec n'importe quelle taille de système n.

3.4 La majorité

3.4.1 Définition

La majorité consiste, à partir d'un ensemble d'agents qui possèdent chacun une donnée d'entrée venant de l'ensemble $\Sigma = \{A, B\}$, à déterminer quel symbole est majoritaire, uniquement en interrogeant un seul agent. En termes plus formels, on considère que n_A (respectivement n_B) est le nombre d'agents possédant le symbole A (respectivement B). Nous appelons ω_A la fonction de sortie qui rend la valeur 1 si le protocole a détecté que les A sont majoritaires, et la valeur 0 s'il a détecté que les B sont majoritaires.

Dans tous les articles que nous avons lus sur le sujet, l'égalité $n_A = n_B$ n'est jamais traitée. Par conséquent, dans ce cas très marginal, nous considérons qu'il est tolérable que le protocole ne converge pas.

Définition 3.4.1 *Soit $\delta \in]0, 1[$. Un protocole de majorité converge à l'instant $\tau \geq 0$ avec la probabilité $1 - \delta$, lorsque, pour tout $t \geq \tau$,*

$$\mathbb{P}\left\{\omega_A\left(C_t^{(i)}\right) = 1_{\{n_A > n_B\}} \text{ pour tout } i \in [\![1, n]\!]\right\} \geq 1 - \delta.$$

Nous pouvons noter que la convergence d'un protocole de majorité correspond à un état du système où tous les nœuds rendent la même valeur, cela s'apparente donc à un protocole de consensus que nous verrons en section 8.1.3.

3.4.2 État de l'art

Le problème de la majorité a été l'un des premiers à avoir été étudié du point de vue du temps de convergence avec un scheduler uniforme. Angluin et al. [13] et, de manière indépendante, Perron et al. [68] proposent un protocole à trois états qui converge avec une probabilité élevée en $O(n \log n)$, mais pas forcément vers la valeur exacte, la valeur exacte n'est garantie que lorsque $|n_A - n_B| = \Omega\left(\sqrt{n \log n}\right)$. De plus, ce protocole tolère une certaine proportion d'agents ($o\left(\sqrt{n}\right)$) au comportement byzantin. Un tel protocole possède un grand intérêt en tant que protocole de consensus (voir section 8.1.3).

Draief et Vojnovic [33] et Mertzios et al. [51] proposent un protocole à quatre états qui résout le problème de la majorité avec un temps parallèle de convergence logarithmique en n mais uniquement au niveau de l'espérance. De plus, l'espérance du temps de convergence tend vers l'infini lorsque γ_A tend vers 1/2.

Alistarh et al. [8] proposent un protocole de population basé sur une méthode de moyenne et conquête ("average and conquer") pour résoudre exactement le problème de la majorité. Leur algorithme utilise deux types d'interactions, la moyenne et la conquête. Le premier type d'interaction est le même que celui utilisé dans nos protocoles de proportion et de comptage (voir chapitre 5). Le second, utilisé pour diffuser le résultat du calcul aux agents dont l'état est zéro, s'apparente plutôt à notre protocole de diffusion de rumeur. Il existe deux versions de ce protocole. L'une, simplifiée pour faire des simulations, permet avec n états de converger vers la majorité avec un temps moyen de $O(\log n)$ et une constante tout à fait raisonnable,

cette version du protocole n'est pas démontrée. L'autre, à laquelle des états sont ajoutés uniquement pour les besoins de la démonstration, converge effectivement en $O(\log n)$ mais avec un constante multiplicative d'au moins 2000, ce qui rend cette version du protocole démontrée peu efficace en simulation.

En 2017 Bilke et al. [22] ont proposé un protocole de majorité qui converge en $O(\log^2 n)$ et nécessite $O(\log^2 n)$ états. Leur protocole est une succession de cycles constitués de deux phases. La première est une phase d'annihilation où, lorsqu'un A rencontre un B, ils se transforment tous les deux en un état neutre Z. La deuxième est une phase de dédoublement où, lorsqu'un Z interagit avec un A (respectivement B), ils deviennent tous les deux A (respectivement B). Le dédoublement s'effectue une fois et une seule pour tous les nœuds non neutres lors d'un cycle. Ensuite, le cycle suivant peut commencer selon le même schéma. Après un maximum de $\log_2 n$ cycles, tous les nœuds ont la même lettre. La principale difficulté de ce protocole est l'orchestration des phases qui est faite, dans ce cas, avec une horloge triviale du type 1-choice. C'est cette horloge qui nécessite tous les états.

En Janvier 2018 Alistarh et al. [6] ont proposé un protocole similaire convergeant en $O(\log^2 n)$ et nécessitant cette fois $O(\log n)$ états. Le principe général est exactement le même que celui du protocole précédent. La seule différence consiste en l'horloge qui dans le cas présent est exactement la même que celle décrite dans le chapitre 6, elle est basée sur le 2-choice.

En Juin 2018 Berenbrick et al. [19] ont proposé, toujours selon le même principe, un protocole résolvant le problème de la majorité avec $O(\log n)$ états et, cette fois, un temps parallèle de convergence de $O(\log^{5/3} n)$. Le gain vient d'une diminution du temps alloué à chaque cycle. Ce temps qui était de $O(\log n)$ dans les protocoles précédents, est maintenant de $O(\log^{2/3} n)$. Les auteurs montrent que les nœuds retardataires, c'est-à-dire qui ne se sont pas dédoublés dans la phase de dédoublement, finissent par pouvoir le faire plus tard sans que cela n'entraîne une augmentation du nombre d'états ou du nombre de cycles. Nous pouvons noter que, curieusement, l'horloge utilisée pour ce protocole est l'horloge triviale correspondant au 1-choice.

3.4.3 Nos résultats

À partir du protocole de comptage défini dans la section 5.2.4, on peut construire un protocole de majorité où la sortie est déduite de la valeur de n_A par rapport à $n/2$. Cela donne un protocole avec $O(n)$ états et un temps de convergence en $O(\log n)$ avec une probabilité élevée. Un protocole de majorité est aussi défini dans la section 5.2 concernant les protocoles de moyenne avec des nombres réels.

3.5 L'horloge globale ou horloge sans leader

3.5.1 Définition

Une horloge globale est un type très particulier de protocole, dans la mesure où il ne converge pas. Il doit toujours rester dans un état où chaque noeud est capable de calculer une estimation du temps parallèle global t/n (le nombre d'interactions ayant eu lieu depuis le démarrage du système divisé par la taille du système) avec

une certaine précision et une probabilité élevée. Cette précision est mesurée par le gap qui est l'écart maximal entre 2 mesures du temps. Ce temps que l'on cherche à mesurer doit se situer entre le minimum et le maximum.

Définition 3.5.1 *Soient* $\delta \in]0;1[$. *Un protocole d'horloge globale donne le temps parallèle avec un gap de* τ *et une probabilité* $1-\delta$ *si pour tout* $t \geq 0$, *sachant que*

$$\begin{aligned} Min(t) &= \min_{1\leq i\leq n} \left\{\omega\left(C_t^{(i)}\right)\right\} \\ Max(t) &= \max_{1\leq i\leq n} \left\{\omega\left(C_t^{(i)}\right)\right\} \\ Gap(t) &= Max(t) - Min(t), \end{aligned}$$

nous avons

$$\frac{t}{n} \in [Min(t), Max(t)] \text{ et } \mathbb{P}\left\{Gap(t) \leq \tau\right\} \geq 1-\delta.$$

3.5.2 État de l'art

Il existe un protocole trivial qui consiste à incrémenter le compteur d'état du noeud à l'initiative d'une interaction. Il s'agit du "1-choice", protocole avec lequel le Gap croît avec la racine carrée du temps parallèle [70], ce qui ne correspond pas à la définition de l'horloge globale. Cependant, grâce à sa simplicité, et dans la mesure où le temps qui doit être mesuré est borné, ce protocole garde un intérêt (voir [22, 19]). La première implémentation d'un dispositif plus complexe a été réalisée en 2018 par Alistarh et al. [6]. Il s'agit d'une astucieuse utilisation de la solution "2-choice" du problème "balls-and-bins" déjà considérablement étudiée dans le domaine de l'équilibrage de charges (load-balancing). Chaque agent possède un compteur initialisé à 0 au démarrage et à chaque interaction, le compteur ayant la valeur la plus basse est incrémenté.

3.5.3 Nos résultats

Le protocole "2-choice" fait l'objet d'une étude approfondie au chapitre 6. L'analyse suit le même cheminement que celui de Peres et al. dans [67], à la différence importante près que nous explicitons et optimisons les constantes.

Chapitre 4

Diffusion de rumeur

Le contexte de ce travail est le domaine bien étudié de la dissémination de l'information dans des systèmes distribués à grande échelle par le biais d'interactions par paires. Ce problème, appelé diffusion de rumeur a été principalement étudié dans le modèle synchrone. Ce modèle repose sur l'hypothèse que tous les nœuds agissent de manière synchrone, c'est-à-dire qu'à chaque ronde du protocole, chaque nœud contacte un voisin aléatoirement, pour échanger leur connaissance de la rumeur, le but étant, qu'à la fin, tout le monde connaisse la rumeur. Dans ce chapitre, nous abandonnons cette hypothèse dans la mesure où elle n'est pas réaliste dans les systèmes à grande échelle. Nous considérons donc la variante asynchrone, où, à des moments aléatoires, les nœuds interagissent successivement par paires, choisies aléatoirement, échangeant leurs informations sur la rumeur. Dans la première partie, nous effectuons une étude du nombre total d'interactions nécessaires pour que tous les nœuds du réseau découvrent la rumeur. La plupart des résultats existants impliquent des constantes non précisées, ce qui ne nous permet pas de comparer différents protocoles. Nous fournissons une analyse approfondie et précise de la distribution de ce nombre total d'interactions, avec son comportement asymptotique. Dans la deuxième partie, nous étendons cette analyse temporelle discrète et considérons le cas du temps continu où un processus de Poisson est associé à chaque nœud pour déterminer les instants auxquels les interactions se produisent. Le temps de propagation de la rumeur est donc plus réaliste car c'est le temps réel nécessaire à tous les nœuds du réseau pour découvrir la rumeur. Nous fournissons une borne fine et un équivalent de la fonction de répartition du temps de propagation de la rumeur. Nous donnons également le comportement asymptotique de la distribution du temps de propagation de la rumeur autour de sa moyenne lorsque le nombre de nœuds tend vers l'infini.

Ce chapitre se base sur trois de nos publications : [57, 60] dans leur intégralité et [59] partiellement.

4.1 Introduction

Dans ce présent chapitre, nous considérons le temps de diffusion de la rumeur dans le modèle push-pull asynchrone. Nous pouvons noter que, en mode asynchrone, lorsque le graphe est complet et que c'est une loi uniforme qui régit les interactions, le mode

push ou le mode pull consiste à ne considérer comme valide qu'une interaction sur deux par rapport au mode push-pull. Par conséquent, il y aura un facteur 2 au niveau de l'espérance, et un facteur 4 au niveau de la variance entre le mode push ou le mode pull d'une part et le mode push-pull d'autre part.

Dans le cas du temps discret, les nœuds interagissent par paires au hasard et si au moins un nœud possède la rumeur, l'autre en est informé. Dans ce cas, le temps de diffusion est défini par le nombre d'interactions nécessaires pour que tous les nœuds du réseau apprennent la rumeur. Dans le cas du temps continu, comme suggéré par Ganesh [37], un processus de Poisson est associé à chaque nœud et lors d'un saut du processus de Poisson d'un nœud, ce nœud choisit de manière aléatoire un autre nœud pour interagir avec lui comme dans le cas discret, c'est-à-dire s'informer de la rumeur si l'un de ces deux nœuds possède la rumeur. Les n processus de Poisson sont supposés être indépendants et de même taux.

Dans un premier temps nous étudions le temps de diffusion de la rumeur dans le modèle push-pull asynchrone à temps discret. Ensuite, nous étendons les résultats obtenus au modèle push-pull asynchrone en temps continu.

Le reste de ce chapitre est organisé comme suit. La section 4.2 présente les résultats obtenus dans [57], il s'agit du modèle à temps discret sur lequel nous nous appuyons pour résoudre le modèle à temps continu. Dans cette section, nous prouvons également la conjecture formulée dans [57]. Plus précisément, si T_n dénote le nombre total d'interactions pour que tous les n nœuds obtiennent la rumeur alors,

$$\lim_{n\longrightarrow\infty} \mathbb{P}\{T_n > \mathbb{E}(T_n)\} \approx 0.448429663727,$$

où $\mathbb{E}(T_n) = (n-1)H_{n-1}$ et H_k est la k-ième somme partielle de la série harmonique.

Dans la section 4.3, nous considérons le modèle en temps continu. Un processus de Poisson est associé à chaque nœud et chaque saut de ces processus de Poisson indépendants correspond à une interaction entre le nœud associé au processus et un autre nœud choisi aléatoirement selon une loi uniforme. Dans ce modèle, le temps nécessaire aux n nœuds pour obtenir la rumeur est Θ_n. Nous donnons d'abord des expressions simples de l'espérance et de la variance de Θ_n. Ensuite, nous donnons une expression explicite de sa distribution et nous obtenons une borne simple de sa fonction complémentaire de répartition. Nous démontrons qu'elle est également un équivalent de sa queue de distribution. On montre aussi que cette borne est beaucoup plus proche de la réalité que les bornes déjà connues. Enfin, nous donnons la distribution limite du rapport $\Theta_n/\mathbb{E}(\Theta_n)$ lorsque le nombre n de nœuds tend vers l'infini. Enfin, la section 4.4 conclut le chapitre.

4.2 Diffusion de rumeur en temps discret

4.2.1 Modélisation

Le modèle est particulièrement simple. Les ensembles d'entrée, de travail et sortie n'ont que 0 et 1 comme éléments, les fonctions d'entrée et de sortie sont l'identité, et la fonction de transition attribue le maximum à chacun des nœuds. Donc en suivant la modélisation de la section 2.2, le protocole est défini par le sextuplet $(\Sigma, \Xi, Q, \iota, \omega, f)$ et on a

- $\Sigma = \Xi = Q = \{0, 1\}$
- $\iota = \omega = Id_{\{0,1\}}$
- $\begin{aligned} \mathrm{f} : \{0,1\} \times \{0,1\} &\longrightarrow \{0,1\} \times \{0,1\} \\ (x, y) &\longmapsto (\max\{x, y\}, \max\{x, y\}) \end{aligned}$

4.2.2 La diffusion

Dans le cas du temps discret, le nombre total d'interactions nécessaires pour que les n nœuds obtiennent la rumeur est noté T_n.

Une valeur 0 ou 1 est associée à chaque nœud. Un nœud avec valeur 1 signifie que ce nœud connaît la rumeur et un nœud avec la valeur 0 signifie qu'il n'est pas au courant de la rumeur. Pour chaque $t \geq 0$, nous désignons par $C_t^{(i)}$ la valeur (0 ou 1) du nœud i à l'instant t.

À chaque instant $t \geq 0$, la configuration t du protocole est notée $C_t = (C_t^{(1)}, \ldots, C_t^{(n)})$, où $C_t^{(i)}$ est l'état (0 ou 1) du nœud i à l'instant t. $C = \{C_t,\ t \geq 0\}$ est un processus stochastique en temps discret sur l'espace d'états $\{0, 1\}^n$.

A chaque instant t, deux indices distincts i et j sont successivement choisis aléatoirement dans l'intervalle $[\![1, n]\!]$. Nous notons X_t la variable aléatoire représentant ce choix à l'instant $t \geq 1$ et nous supposons que ce choix est uniforme, c'est-à-dire que nous supposons que

$$\mathbb{P}\{X_t = (i, j)\} = \frac{1}{n(n-1)} 1_{\{i \neq j\}}, \quad \forall i, j \in [\![1, n]\!].$$

Une fois le couple (i, j) choisi à l'instant $t \geq 1$, la manière dont le processus atteint l'état C_t est donnée par

$$C_t^{(i)} = C_t^{(j)} = \max\{C_{t-1}^{(i)}, C_{t-1}^{(j)}\} \text{ et } C_t^{(r)} = C_{t-1}^{(r)} \text{ pour } r \neq i, j.$$

La variable aléatoire T_n, définie par

$$T_n = \inf \left\{ t \geq 0 \,\middle|\, \sum_{i=1}^{n} C_t^{(i)} = n \right\},$$

représente le nombre d'interactions nécessaires pour que tous les nœuds du réseau connaissent la rumeur.

Nous introduisons le processus stochastique en temps discret $Y = \{Y_t,\ t \geq 0\}$ avec comme ensemble d'états $[\![1, n]\!]$ défini, pour tout $t \geq 0$, par

$$Y_t = \sum_{i=1}^{n} C_t^{(i)}.$$

La variable aléatoire Y_t représente le nombre de nœuds connaissant la rumeur à l'instant t. Le processus stochastique Y est une chaîne de Markov homogène à n

états, de comme matrice des probabilités de transition notée A. Les probabilités de transition différentes de zéro sont données pour $i, j \in [\![1, n]\!]$ par

$$\begin{cases} A_{i,i} &= 1 - \dfrac{2i(n-i)}{n(n-1)}, \\ \\ A_{i,i+1} &= \dfrac{2i(n-i)}{n(n-1)}, \text{ pour } i \neq n. \end{cases}$$

En effet, quand $Y_t = i$, pour obtenir $Y_{t+1} = i + 1$, soit le premier nœud doit être choisi parmi ceux ayant l'état 1 (probabilité i/n) et le second agent doit être choisi parmi ceux ayant l'état 0 (probabilité $(n-i)/(n-1)$) ou le premier nœud doit être choisi parmi ceux ayant l'état 0 (probabilité $(n-i)/n$) et le deuxième nœud doit être choisi parmi ceux ayant l'état 1 (probabilité $i/(n-1)$), les états $1, \ldots, n-1$ étant des états transitoires et l'état n étant l'état absorbant. Pour $i \in [\![1, n-1]\!]$, nous introduisons la notation

$$p_i = \mathbb{P}\{Y_{t+1} = i + 1 \mid Y_t = i\} = \frac{2i(n-i)}{n(n-1)}.$$

Avec cette notation la figure 4.1 représente le graphe des transitions de Y.

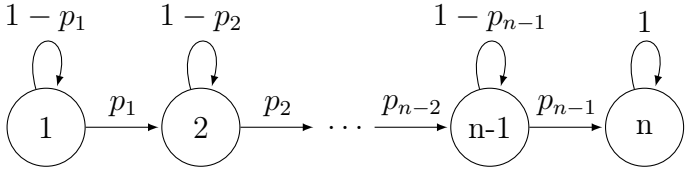


FIGURE 4.1 : Graphe des transitions de la chaîne de Markov Y.

La variable aléatoire T_n peut s'écrire

$$T_n = \inf\{t \geq 0 \mid Y_t = n\}.$$

Il est bien connu, voir par exemple [72], que la distribution de T_n est donnée, pour chaque $k \geq 0$, par

$$\mathbb{P}\{T_n > k\} = \alpha Q^k \mathbb{1}, \tag{4.1}$$

où α est le vecteur ligne contenant les probabilités initiales des états $1, \ldots, n-1$, c'est-à-dire $\alpha_i = \mathbb{P}\{Y_0 = i\} = 1_{\{i=1\}}$, Q est la sous-matrice obtenue à partir de A en supprimant la ligne et la colonne correspondant à l'état absorbant n et $\mathbb{1}$ est le vecteur colonne de dimension $n-1$ dont toutes les composantes sont égales à 1.

Q est la matrice contenant les probabilités de transition entre les différents états transitoires c'est-à-dire

$$\begin{cases} Q_{i,i} &= 1 - p_i \quad \text{pour } i \in [\![1, n-1]\!] \\ Q_{i,i+1} &= p_i \quad \text{pour } i \in [\![1, n-2]\!]. \end{cases} \tag{4.2}$$

Nous appelons H_k la $k-$ième somme partielle de la série harmonique, c'est-à-dire

$$H_0 = 0 \text{ et } H_k = \sum_{\ell=1}^{k} \frac{1}{\ell}, \text{ pour } k \geq 1.$$

Soit S_i, pour $i \in [\![1, n-1]\!]$, le temps passé par la chaîne de Markov Y dans l'état i, ou, autrement dit, le nombre d'interactions pendant lequel le nombre de nœuds connaissant la rumeur reste égal à i. S_i suit la loi géométrique de paramètre p_i et nous avons

$$\text{si } Y_0 = i \text{ alors } T_n = \sum_{\ell=i}^{n-1} S_\ell.$$

4.2.3 Analyse du temps de diffusion

Dans ce qui suit nous étudions l'espérance et la variance de T_n, le nombre d'interactions nécessaires pour que tous les nœuds du réseau connaissent la rumeur. Nous fournissons ensuite une expression explicite de la distribution de T_n, puis une borne et un équivalent pour la queue de la distribution de T_n.

4.2.3.1 Espérance et variance de T_n

L'espérance du temps nécessaire pour que tous les nœuds connaissent la rumeur $\mathbb{E}(T_n)$ est donnée par

$$\mathbb{E}(T_n) = \alpha(I - Q)^{-1}\mathbb{1}, \tag{4.3}$$

ou I est la matrice identité. Cette espérance peut aussi s'écrite

$$\mathbb{E}(T_n) = \sum_{i=1}^{n-1} \alpha_i \mathbb{E}(T_n \mid Y_0 = i).$$

Elle est donnée par le théorème suivant.

Théorème 4.2.1 *Pour tout $n \geq 1$ et $i \in \{1, \ldots, n\}$, nous avons*

$$\mathbb{E}(T_n \mid Y_0 = i) = \frac{(n-1)(H_{n-1} + H_{n-i} - H_{i-1})}{2}.$$

Preuve. Si $Y_0 = n$, cela veut dire que tous les nœuds commencent en connaissant la rumeur, donc nous avons $T_n = 0$ et aussi $\mathbb{E}(T_n \mid Y_0 = n) = 0$.

Pour $i \in [\![1, n-1]\!]$ nous avons, en rappelant que $\mathbb{E}(S_\ell) = 1/p_\ell$,

$$\begin{aligned}
\mathbb{E}(T_n \mid Y_0 = i) &= \sum_{\ell=i}^{n-1} \mathbb{E}(S_\ell) \\
&= \sum_{\ell=i}^{n-1} \frac{1}{p_\ell} \\
&= \frac{n(n-1)}{2} \sum_{\ell=i}^{n-1} \frac{1}{\ell(n-\ell)} \\
&= \frac{n-1}{2} \sum_{\ell=i}^{n-1} \left(\frac{1}{\ell} + \frac{1}{n-\ell} \right) \\
&= \frac{(n-1)(H_{n-1} + H_{n-i} - H_{i-1})}{2},
\end{aligned}$$

ce qu'il fallait démonter. ■

En particulier quand l'initiateur de la rumeur est unique, c'est-à-dire quand $Y_0 = 1$, nous avons

$$\mathbb{E}(T_n) = \mathbb{E}(T_n \mid Y_0 = 1) = (n-1)H_{n-1} \underset{n\longrightarrow\infty}{\sim} n\ln(n).$$

La variance (que nous notons $\mathbb{V}$) de T_n est obtenue de manière similaire par une sommation de résultats connus sur la loi géométrique.

Théorème 4.2.2 *Pour tout $n \geq 1$ et $i \in [\![1, n]\!]$, nous avons*

$$\mathbb{V}(T_n \mid Y_0 = i) = \frac{(n-1)^2}{4}\left(\sum_{\ell=i}^{n-1}\frac{1}{\ell^2} + \sum_{\ell=1}^{n-i}\frac{1}{\ell^2}\right) - \frac{\mathbb{E}(T_n \mid Y_0 = i)}{n}.$$

Preuve. Si $Y_0 = n$, cela veut dire que tous les nœuds commencent en connaissant la rumeur donc dans ce cas $T_n = 0$ et par conséquent $\mathbb{V}(T_n \mid Y_0 = n) = 0$. Pour $i \in \{1, \ldots, n-1\}$ nous avons, en utilisant l'indépendance des S_ℓ,

$$\begin{aligned}
\mathbb{V}(T_n \mid Y_0 = i) &= \sum_{\ell=i}^{n-1}\mathbb{V}(S_\ell) = \sum_{\ell=i}^{n-1}\frac{1-p_\ell}{p_\ell^2} = \sum_{\ell=i}^{n-1}\frac{1}{p_\ell^2} - \sum_{\ell=i}^{n-1}\frac{1}{p_\ell} \\
&= \frac{n^2(n-1)^2}{4}\sum_{\ell=i}^{n-1}\frac{1}{\ell^2(n-\ell)^2} - \frac{n(n-1)}{2}\sum_{\ell=i}^{n-1}\frac{1}{\ell(n-\ell)} \\
&= \frac{(n-1)^2}{4}\sum_{\ell=i}^{n-1}\left(\frac{1}{\ell} + \frac{1}{n-\ell}\right)^2 - \frac{n(n-1)}{2}\sum_{\ell=i}^{n-1}\frac{1}{\ell(n-\ell)} \\
&= \frac{(n-1)^2}{4}\sum_{\ell=i}^{n-1}\left(\frac{1}{\ell^2} + \frac{1}{(n-\ell)^2}\right) - \frac{n-1}{2}\sum_{\ell=i}^{n-1}\frac{1}{\ell(n-\ell)} \\
&= \frac{(n-1)^2}{4}\sum_{\ell=i}^{n-1}\left(\frac{1}{\ell^2} + \frac{1}{(n-\ell)^2}\right) - \frac{\mathbb{E}(T_n \mid Y_0 = i)}{n} \\
&= \frac{(n-1)^2}{4}\left(\sum_{\ell=i}^{n-1}\frac{1}{\ell^2} + \sum_{\ell=1}^{n-i}\frac{1}{\ell^2}\right) - \frac{\mathbb{E}(T_n \mid Y_0 = i)}{n},
\end{aligned}$$

ce qu'il fallait démontrer. ∎

En particulier, quand l'initiateur de la rumeur est unique, c'est-à-dire quand $Y_0 = 1$, nous avons

$$\begin{aligned}
\mathbb{V}(T_n) &= \mathbb{V}(T_n \mid Y_0 = 1) \\
&= \frac{(n-1)^2}{2}\sum_{\ell=1}^{n-1}\frac{1}{\ell^2} - \frac{n-1}{n}H_{n-1} \underset{n\longrightarrow\infty}{\sim} \frac{\pi^2 n^2}{12}.
\end{aligned}$$

Plus généralement, à partir du théorème 4.2.2, nous avons

$$\begin{aligned}
\mathbb{V}(T_n \mid Y_0 = i) &\leq \frac{(n-1)^2}{4}\left(\sum_{\ell=i}^{n-1}\frac{1}{\ell^2} + \sum_{\ell=1}^{n-i}\frac{1}{\ell^2}\right) \\
&\leq \frac{(n-1)^2}{2}\sum_{\ell=1}^{n-1}\frac{1}{\ell^2} \leq \frac{\pi^2 n^2}{12}.
\end{aligned}$$

Il s'ensuit que

$$\mathbb{V}(T_n) = \sum_{i=1}^{n-1} \alpha_i \mathbb{V}(T_n \mid Y_0 = i) \leq \frac{\pi^2 n^2}{12}.$$

4.2.3.2 Expression explicite de la distribution de T_n

La distribution de T_n, pour $n \geq 2$, est donnée par la relation (4.1). Elle peut être explicitée comme suit. Soit $V(k) = (V_1(k), \ldots, V_{n-1}(k))$ le vecteur colonne défini par $V_i(k) = \mathbb{P}\{T_n > k \mid Y_0 = i\}$. En se basant sur la relation (4.1), nous avons $V(k) = Q^k \mathbb{1}$. Comme $V(0) = \mathbb{1}$, en écrivant $V(k) = QV(k-1)$ pour $k \geq 1$, nous obtenons

$$\begin{cases} V_i(k) & = (1-p_i)\, V_i(k-1) + p_i V_{i+1}(k-1), \text{ pour } i \in [\![1, n-2]\!], \\ V_{n-1}(k) & = (1-p_{n-1})\, V_{n-1}(k-1). \end{cases} \tag{4.4}$$

Rappelons-nous que nous avons $p_i = 2i(n-i)/(n(n-1))$. Cette récurrence peut être explicitée car nous avons, pour $k \geq 0$,

$$V_{n-1}(k) = (1-p_{n-1})^k = \left(1 - \frac{2}{n}\right)^k. \tag{4.5}$$

Dans le théorème suivant, nous déduisons de la récurrence (4.4), une formule explicite de la distribution de T_n.

Théorème 4.2.3 *Pour tout $n \geq 1$, $k \geq 0$ et $i \in [\![1, n-1]\!]$, nous avons*

$$\mathbb{P}\{T_n > k \mid Y_0 = n - i\} = \sum_{j=1}^{\lfloor n/2 \rfloor} \left(c_{i,j}(1-p_j) + k d_{i,j}\right)(1-p_j)^{k-1},$$

où les coefficients $c_{i,j}$ et $d_{i,j}$, qui ne dépendent pas de k, sont donnés, pour $j \in [\![1, n-1]\!]$, par

$$c_{1,j} = 1_{\{j=1\}} \text{ et } d_{1,j} = 0$$

et pour $(i,j) \in [\![2, n-1]\!] \times [\![1, n-1]\!]$, par

$$\begin{cases} c_{i,j} & = \dfrac{p_i c_{i-1,j}}{p_i - p_j} - \dfrac{p_i d_{i-1,j}}{(p_i - p_j)^2} & \text{pour} & i \neq j, n-j, \\ d_{i,j} & = \dfrac{p_i d_{i-1,j}}{p_i - p_j} & \text{pour} & i \neq j, n-j, \\ c_{i,i} & = 1 - \displaystyle\sum_{j=1, j\neq i}^{\lfloor n/2 \rfloor} c_{i,j} & \text{pour} & i \leq \lfloor n/2 \rfloor, \\ c_{i,n-i} & = 1 - \displaystyle\sum_{j=1, j\neq n-i}^{\lfloor n/2 \rfloor} c_{i,j} & \text{pour} & i > \lfloor n/2 \rfloor, \\ d_{i,i} & = p_i c_{i-1,i} & \text{pour} & i \leq \lfloor n/2 \rfloor, \\ d_{i,n-i} & = p_i c_{i-1,n-i} & \text{pour} & i > \lfloor n/2 \rfloor. \end{cases} \tag{4.6}$$

Preuve. Voir annexe section A.1. ■

4.2.3.3 Bornes et queue de la distribution de T_n

La formule exacte de la distribution de T_n présentée plus haut est assez complexe à utiliser en pratique, et le calcul des coefficients du théorème 4.2.3 peut prendre beaucoup de temps pour de grandes valeurs de n. Pour simplifier ce problème, nous proposons dans le théorème suivant un majorant qui est aussi un équivalent de la quantité $\mathbb{P}\{T_n > k \mid Y_0 = i\}$. Ces calculs sont déduits de la formule de récurrence (4.4).

Théorème 4.2.4 *Pour tout $n \geq 2$ et $k \geq 1$, nous avons*

$$\mathbb{P}\{T_n > k \mid Y_0 = 1\} \leq \left(1 + \frac{2k(n-2)^2}{n}\right)\left(1 - \frac{2}{n}\right)^{k-1},$$

$$\mathbb{P}\{T_n > k \mid Y_0 = 1\} \underset{k\longrightarrow\infty}{\sim} \left(1 + \frac{2k(n-2)^2}{n}\right)\left(1 - \frac{2}{n}\right)^{k-1}$$

et pour $i \in [\![2, n-1]\!]$ et $k \geq 0$,

$$\mathbb{P}\{T_n > k \mid Y_0 = i\} \leq \frac{(n-i)(n-2)}{i-1}\left(1 - \frac{2}{n}\right)^{k},$$

$$\mathbb{P}\{T_n > k \mid Y_0 = i\} \underset{k\longrightarrow\infty}{\sim} \frac{(n-i)(n-2)}{i-1}\left(1 - \frac{2}{n}\right)^{k}.$$

D'autre part, nous avons

$$\mathbb{P}\{T_n > k\} \leq \mathbb{P}\{T_n > k \mid Y_0 = 1\}.$$

Preuve. Voir annexe section A.1. ■

Le théorème précédent fournit deux formules : une pour le cas où un seul nœud connaît la rumeur au départ, une autre pour le cas où deux nœuds ou plus connaissent la rumeur au départ. Cette dernière amène simplement à une évaluation très précise du temps de convergence avec probabilité élevée, à partir de deux nœuds connaissant la rumeur.

Théorème 4.2.5 *Pour tout $\delta \in]0; 1[$, nous avons*

$$\mathbb{P}\left\{T_n \leq \lceil n\left(\ln(n) - \ln(\delta)/2\right)\rceil \mid Y_0 = 2\right\} \geq 1 - \delta.$$

Preuve. En appliquant le théorème 4.2.4 avec $i = 2$ nous obtenons, pour tout $k \geq 0$

$$\mathbb{P}\{T_n > k \mid Y_0 = 2\} \leq (n-2)^2\left(1 - \frac{2}{n}\right)^{k}.$$

En prenant $k = \lceil n\left(\ln(n) - \ln(\delta)/2\right)\rceil$, nous pouvons écrire

$$\left(1 - \frac{2}{n}\right)^{\lceil n(\ln(n)-\ln(\delta)/2)\rceil} \leq \left(1 - \frac{2}{n}\right)^{n(\ln(n)-\ln(\delta)/2)} = e^{n(\ln(n)-\ln(\delta)/2)\ln(1-2/n)}.$$

En utilisant le fait que $\ln(1-x) \leq -x$, pour tout $x \in [0,1[$, nous avons

$$\left(1-\frac{2}{n}\right)^{\lceil n(\ln(n)-\ln(\delta)/2)\rceil} \leq e^{-2\ln(n)+\ln(\delta)} = \frac{\delta}{n^2}$$

et par conséquent

$$\mathbb{P}\{T_n > \lceil n\,(\ln(n)-\ln(\delta)/2)\rceil \mid Y_0 = 2\} \leq \frac{(n-2)^2\delta}{n^2} \leq \delta,$$

donc

$$\mathbb{P}\{T_n \leq \lceil n\,(\ln(n)-\ln(\delta)/2)\rceil \mid Y_0 = 2\} \geq 1-\delta,$$

ce qu'il fallait démontrer. ■

Notons que la preuve de ce lemme n'utilise pas l'inégalité de Markov. Les approximations ont toutes été faites avec des équivalents, ce qui signifie que le résultat de ce théorème est très proche de la réalité. Cela sera illustré dans la section 4.2.5. Ce dernier théorème possède un grand avantage pratique, il a été utilisé dans le chapitre 7 et dans l'article [59].

La limite établie dans le théorème 4.2.4 est d'autant plus intéressante que les inégalités probabilistes habituelles ne parviennent pas à fournir des résultats pertinents dans ce cas particulier. Par exemple, l'inégalité de Markov conduit à ce que pour tous les nombres réels $c \geq 1$

$$\mathbb{P}\{T_n \geq c\mathbb{E}(T_n)\} \leq \frac{1}{c},$$

et l'inégalité de Bienaymé-Tchebychev mène à ce que pour tous les nombres réels $x > 0$

$$\mathbb{P}\{|T_n - \mathbb{E}(T_n)| \geq x\} \leq \frac{\pi^2 n^2}{12x^2}.$$

Janson [44] fournit une borne, basée sur l'inégalité de Chernoff, pour la queue de la distribution d'une somme de variables aléatoires de loi géométrique, indépendantes mais non nécessairement identiquement distribuées. Dans le cas particulier de notre protocole de diffusion de rumeur, cela mène au résultat qui suit.

Théorème 4.2.6 *Pour tout $n \geq 3$ et pour tout nombre réel $c \geq 1$, nous avons*

$$\mathbb{P}\left\{T_n > c\mathbb{E}(T_n)\right\} \leq \frac{1}{c}\left(1-\frac{2}{n}\right)^{(c-1-\ln c)(n-1)H_{n-1}}.$$

Le terme de droite est égal à 1 quand $c = 1$.

Preuve. Nous avons déjà vu que

$$\mathbb{P}\left\{T_n > c\mathbb{E}(T_n)\right\} \leq \mathbb{P}\{T_n > c\mathbb{E}(T_n) \mid Y_0 = 1\}.$$

La borne supérieure est une application du théorème 2.3 de [44], et elle est clairement égale à 1 quand $c = 1$. ■

En appliquant le théorème 4.2.4 pour $k = \lfloor c\mathbb{E}(T_n)\rfloor$, nous obtenons

$$\begin{aligned}\mathbb{P}\{T_n > c\mathbb{E}(T_n)\} &\leq \left(1 + \frac{2\lfloor c\mathbb{E}(T_n)\rfloor(n-2)^2}{n}\right) \times \left(1 - \frac{2}{n}\right)^{\lfloor c\mathbb{E}(T_n)\rfloor - 1} \\ &\leq \left(1 + \frac{2c\mathbb{E}(T_n)(n-2)^2}{n}\right) \times \left(1 - \frac{2}{n}\right)^{c\mathbb{E}(T_n)-2}.\end{aligned}$$

A partir de maintenant, nous notons cette borne $f(c,n)$ et, de manière similaire, nous notons $g(c,n)$ la borne de $\mathbb{P}(T_n > c\mathbb{E}(T_n))$ dérivée du théorème 4.2.6. Nous obtenons donc, pour $n \geq 3$ et $c \geq 1$,

$$\begin{aligned}f(c,n) &= \left(1 + \frac{2c(n-1)H_{n-1}(n-2)^2}{n}\right)\left(1 - \frac{2}{n}\right)^{c(n-1)H_{n-1}-2} \\ g(c,n) &= \frac{1}{c}\left(1 - \frac{2}{n}\right)^{(c-1-\ln(c))(n-1)H_{n-1}}.\end{aligned}$$

Nous introduisons également la notation

$$e(c,n) = \mathbb{P}\{T_n > c\mathbb{E}(T_n)\}.$$

Théorème 4.2.7 *Pour tout $n \geq 3$, il existe un unique $c^* \geq 1$ tel que $f(c^*,n) = g(c^*,n)$ et nous avons*

$$\begin{cases} f(c,n) > g(c,n) & \text{pour tout } 1 \leq c < c^* \\ \\ f(c,n) < g(c,n) & \text{pour tout } c > c^*. \end{cases} \tag{4.7}$$

De plus,

$$\lim_{c \longrightarrow \infty} \frac{f(c,n)}{g(c,n)} = 0.$$

Preuve. Voir annexe section A.1 ■

La valeur c^* que nous avons introduite dans le théorème précédent, correspond à la borne à partir de laquelle la fonction $f(c,n)$ issue de notre étude est une meilleure approximation de $e(c,n) = \mathbb{P}\{T_n > c\mathbb{E}(T_n)\}$ que la fonction $g(c,n)$ issue des travaux de Janson [44].

Les graphes des figures 4.2, 4.3 et 4.4 illustrent le comportement des bornes $f(c,n)$ et $g(c,n)$, en fonction de c et pour différentes valeurs de n, comparé à la distribution exacte de T_n au point $c\mathbb{E}(T_n)$, c'est-à-dire par rapport à $e(c,n) = \mathbb{P}\{T_n > \mathbb{E}(T_n)\}$.

n	10	10^2	10^3	10^4	10^5	10^6	10^7
c^*	1.09	1.15	1.13	1.11	1.10	1.09	1.08

Tableau 4.1 : Valeurs approximatives de c^* pour différentes valeurs de la taille du système n.

La borne $f(c,n)$, qui provient du théorème 4.2.4, se montre clairement meilleure que la borne de Chernoff $g(c,n)$ provenant de [44] au delà du seuil c^* défini au théorème 4.2.7. De plus, ce seuil semble tendre vers 1 quand n tend vers l'infini, comme nous pouvons le voir dans le tableau 4.1.

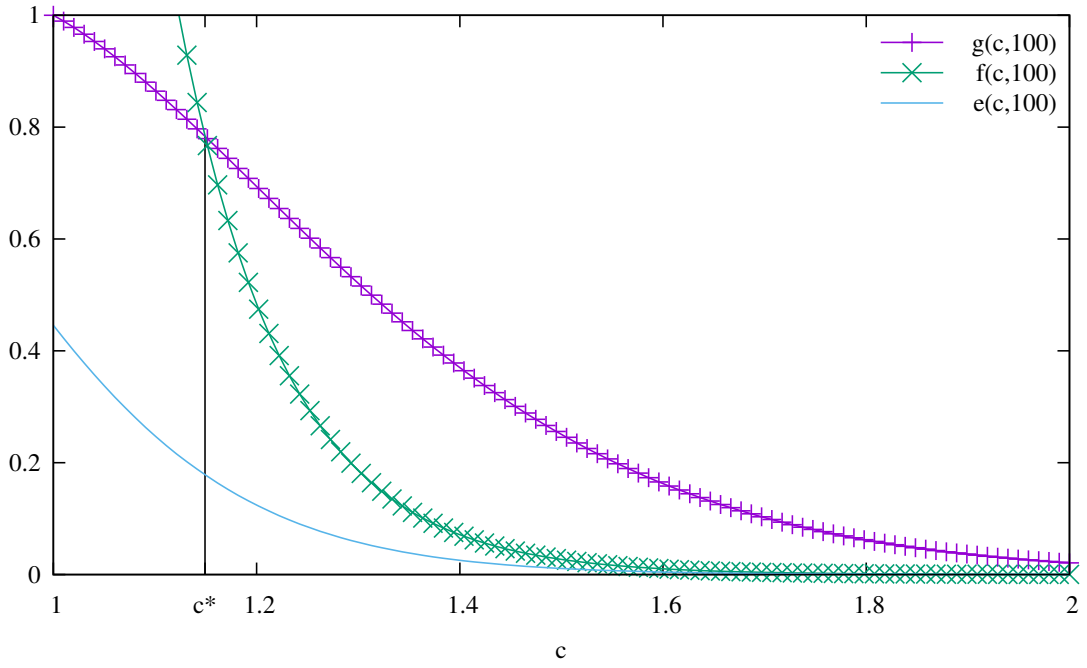


FIGURE 4.2 : Bornes $f(c,n)$ et $g(c,n)$ comparées à la valeur exacte de $\mathbb{P}(T_n > c\mathbb{E}(T_n)) = e(c,n)$ pour $n = 100$, en fonction de c. Dans ce cas, nous avons $c^* = 1.14641$.

4.2.4 Analyse asymptotique de la distribution de T_n

Nous analysons, dans cette section, le comportement de la distribution de T_n au point $c\mathbb{E}(T_n)$ lorsque n tend vers l'infini, en fonction de la valeur de c.

Nous montrons dans le corollaire suivant que les bornes $f(c,n)$ et $g(c,n)$, obtenues à partir du théorème 4.2.4 et du théorème 4.2.6 respectivement avec $k = c\mathbb{E}(T_n)$, tendent toutes les deux vers 0 lorsque n tend vers l'infini.

Corollaire 4.2.8 *Pour tout nombre réel $c > 1$, nous avons*

$$\lim_{n\longrightarrow\infty} f(c,n) = 0 \text{ et } \lim_{n\longrightarrow\infty} g(c,n) = 0.$$

Preuve. Pour tout $x \in [0,1)$, nous avons $\ln(1-x) \leq -x$. En appliquant cette propriété à la borne $f(c,n)$ nous obtenons

$$\begin{aligned} f(c,n) &\leq \left(1 + \tfrac{2c(n-1)H_{n-1}(n-2)^2}{n}\right) e^{-2(c(n-1)H_{n-1}-2)/n} \\ &\leq (1 + 2c(n-2)^2 H_{n-1})\, e^{-2(c(n-1)H_{n-1}-2)/n}. \end{aligned}$$

Comme $\ln(n) \leq H_{n-1} \leq 1 + \ln(n-1)$, nous avons

$$\begin{aligned} f(c,n) &\leq \left(1 + 2c(n-2)^2(1+\ln(n-1)\right) e^{-2(c(n-1)\ln(n)-2)/n} \\ &= \left(1 + 2c(n-2)^2(1+\ln(n-1))\right) e^{-2c\ln(n)} e^{2(c\ln(n)+2)/n} \\ &= \frac{1 + 2c(n-2)^2(1+\ln(n-1))}{n^{2c}} e^{2(c\ln(n)+2)/n}. \end{aligned}$$

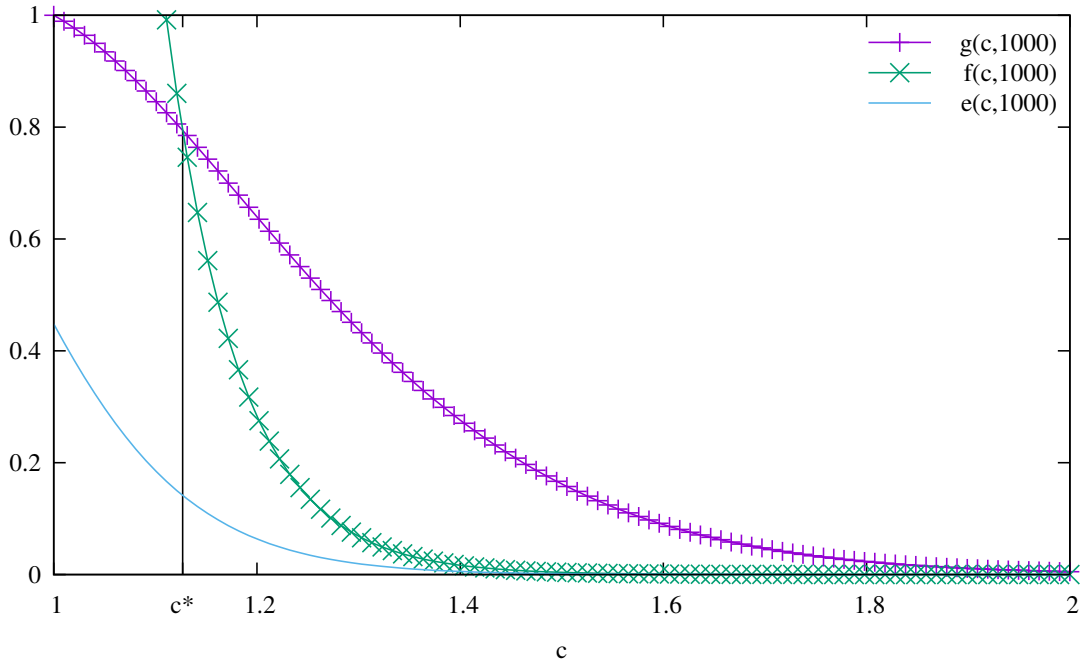


FIGURE 4.3 : Bornes $f(c,n)$ et $g(c,n)$ comparées à la valeur exacte de $\mathbb{P}(T_n > c\mathbb{E}(T_n)) = e(c,n)$ pour $n = 1000$, en fonction de c. Dans ce cas, nous avons $c^* = 1.12673$.

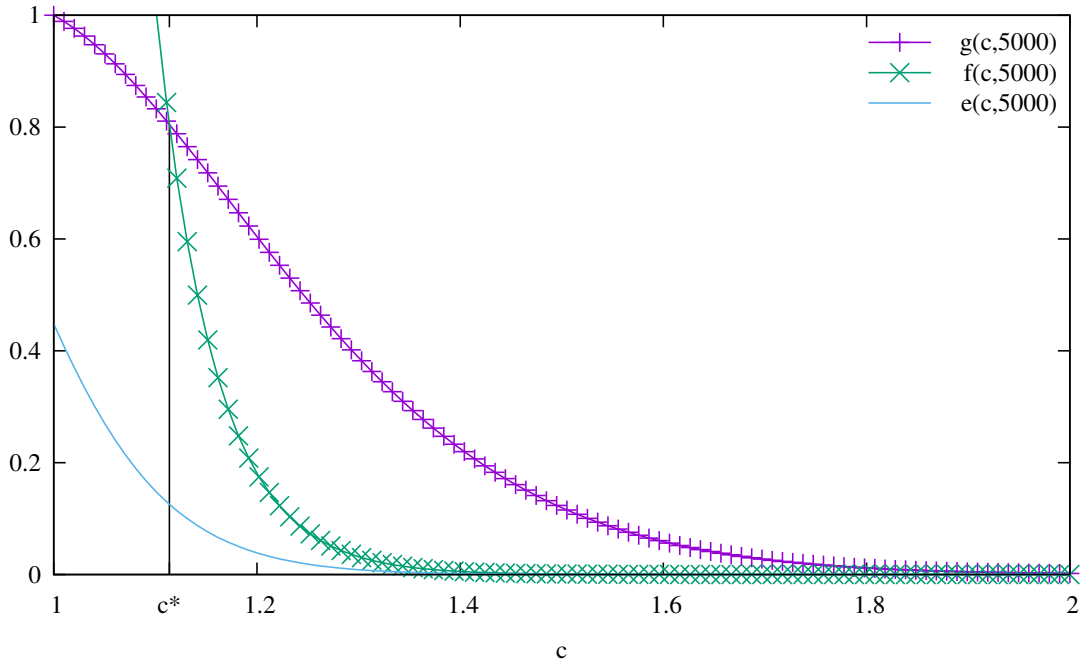


FIGURE 4.4 : Bornes $f(c,n)$ et $g(c,n)$ comparées à la valeur exacte de $\mathbb{P}(T_n > c\mathbb{E}(T_n)) = e(c,n)$ pour $n = 5000$, en fonction de c. Dans ce cas, nous avons $c^* = 1.11385$.

Pour $x \geq 0$, la fonction $u(x) = e^{2(c\ln(x)+2)/x}$ vérifie $\lim_{x\longrightarrow\infty} u(x) = 1$. De même, comme $c > 1$, la fonction $v(x) = (1+2c(x-2)^2(1+\ln(x-1))/x^{2c}$ vérifie $\lim_{x\longrightarrow\infty} v(x) = 0$, ce qui implique que $\lim_{n\longrightarrow\infty} f(c,n) = 0$.

En ce qui concerne la borne $g(c,n)$, nous avons

$$\begin{aligned} g(c,n) &= \frac{1}{c}\left(1-\frac{2}{n}\right)^{(c-1-\ln(c))(n-1)H_{n-1}} \\ &= \frac{1}{c}e^{(c-1-\ln(c))(n-1)H_{n-1}\ln(1-2/n)} \\ &\leq \frac{1}{c}e^{-2(c-1-\ln(c))(n-1)H_{n-1}/n}, \end{aligned}$$

qui tend vers 0 quand n tend vers l'infini, car $c-1-\ln(c)$ est strictement positif pour $c > 1$. ■

Théorème 4.2.9 *Pour tout $c \neq 1$, nous avons*

$$\lim_{n\to+\infty} \mathbb{P}\{T_n > c\mathbb{E}(T_n)\} = \begin{cases} 0 & si\ c > 1 \\ \\ 1 & si\ c < 1. \end{cases}$$

Preuve. A partir du corollaire 4.2.8, les deux bornes $f(c,n)$ et $g(c,n)$ de $\mathbb{P}\{T_n > c\mathbb{E}(T_n)\}$ tendent vers 0 quand n tend vers l'infini. Aussi, en utilisant $f(c,n)$ ou $g(c,n)$, nous obtenons

$$\lim_{n\longrightarrow\infty} \mathbb{P}\{T_n > c\mathbb{E}(T_n)\} = 0 \text{ pour tout } c > 1.$$

Dans le cas où $c < 1$, le théorème 3.1 de [44] mène à

$$\mathbb{P}\{T_n > c\mathbb{E}(T_n)\} \geq 1 - e^{-2(n-1)H_{n-1}(c-1-\ln(c))/n} \geq 1 - e^{-2(n-1)\ln(n)(c-1-\ln(c))/n}.$$

Comme $c-1-\ln(c) > 0$ pour tout $c \in [0,1)$, le terme de droite de cette inégalité tend vers 1 quand $n \longrightarrow \infty$. Par conséquent, $\lim_{n\longrightarrow\infty} \mathbb{P}\{T_n > c\mathbb{E}(T_n)\} = 1$ quand $c < 1$. ■

Pour $c = 1$, ce résultat avait été formulé dans [57] comme une conjecture. Nous sommes maintenant capables d'en donner une preuve que nous avons publiée dans [60].

Théorème 4.2.10

$$\lim_{n\longrightarrow\infty} \mathbb{P}\{T_n > \mathbb{E}(T_n)\} = 1 - 2e^{-\gamma}K_1\left(2e^{-\gamma}\right) \approx 0.448429663727,$$

où γ est la constante d'Euler donnée par $\gamma = \lim_{n\longrightarrow\infty}(H_n - \ln(n)) \approx 0.5772156649$ et K_1 est la fonction de Bessel modifiée de deuxième espèce et d'ordre 1, donnée pour $z > 0$, par

$$K_1(z) = \frac{z}{4}\int_0^{+\infty} t^{-2}e^{-t-z^2/4t}dt.$$

Preuve. Voir annexe section A.1 ■

4.2.5 Simulation de la diffusion de rumeur

Cette section montre à quel point notre limite donnée dans le théorème 4.2.5 est proche des résultats de simulation.

Une simulation est constituée des étapes suivantes : premièrement, les états des n nœuds sont initialisés à 0 à l'exception de deux nœuds dont les états sont initialisés à 1. Ensuite, à chaque étape de la simulation, deux nœuds sont choisis aléatoirement selon une loi uniforme pour interagir et mettre à jour leur état, en conservant la valeur maximale de leurs états. La simulation s'arrête lorsque les états de tous les nœuds sont égaux à 1. Nous avons exécuté N simulations indépendantes. Pour chacune, nous avons stocké et ordonné les N valeurs des temps de convergence indiqués par $\theta_1 \leq \ldots \leq \theta_N$. Rappelons que le temps de convergence θ_i est le nombre d'interactions qui ont été nécessaires pour que tous les nœuds du système aient leur état égal à 1. L'estimation de l'instant τ tel que $\mathbb{P}\{T_n < n\tau\} \geq 1 - \delta$ est donc donnée par la valeur $\theta_{\lceil N(1-\delta)\rceil}$.

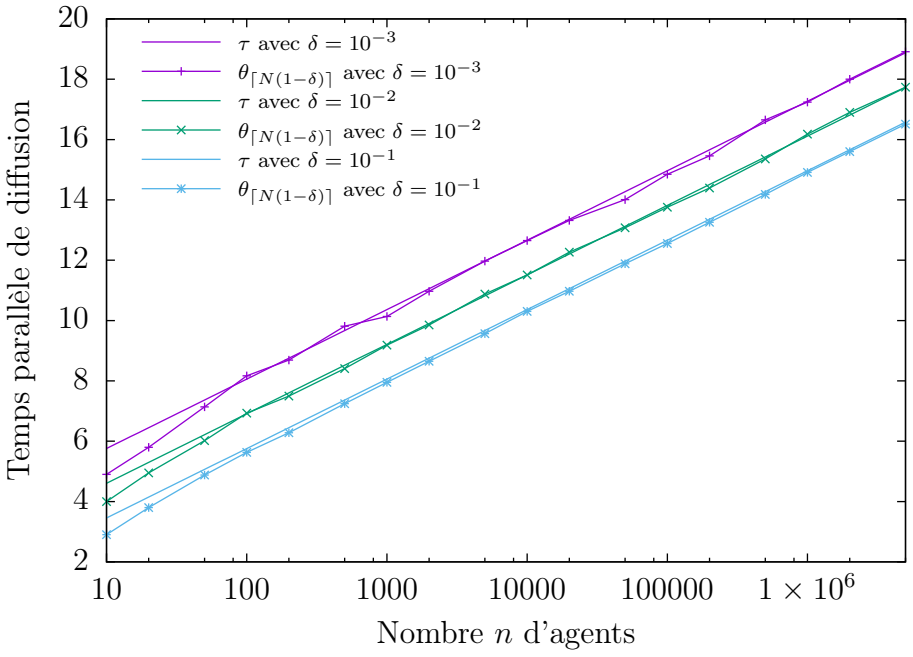


FIGURE 4.5 : Temps parallèle de convergence de la diffusion de rumeur en fonction de n, avec $N = 10^4$.

Rappelons que le temps de convergence parallèle est égal au temps de convergence divisé par n. Les figures 4.5 et 4.6 représentent les temps parallèles de convergence $\theta_{\lceil N(1-\delta)\rceil}/n$ et $\tau = \ln n + 0.5 \ln \delta$, pour différentes valeurs de δ pour la première, et pour différentes valeurs n pour la seconde. Les deux figures montrent que les résultats théoriques sont très proches des résultats des simulations.

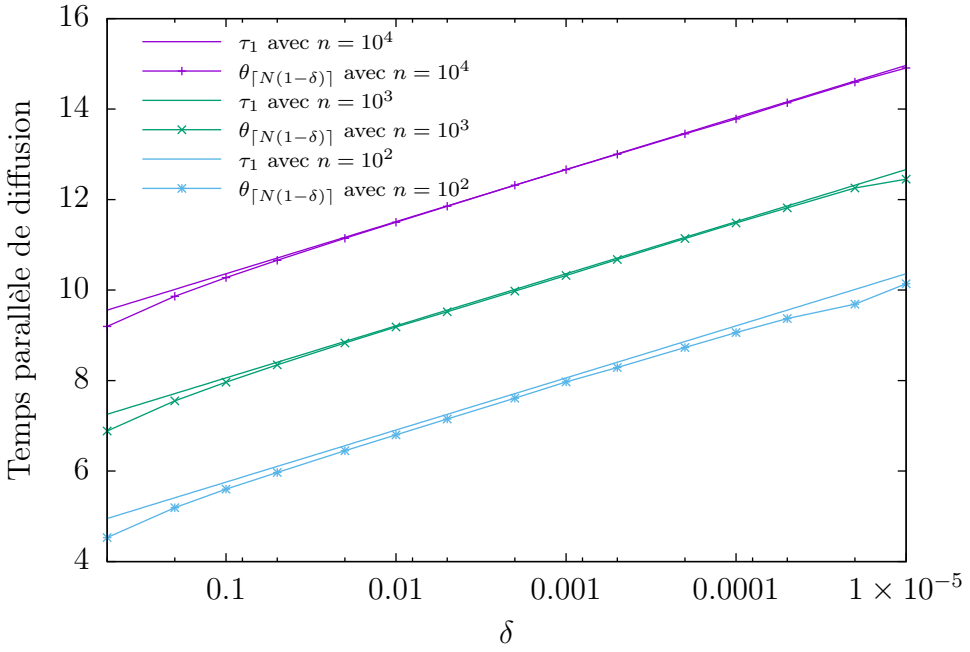


FIGURE 4.6 : Temps parallèle de convergence de la diffusion de rumeur en fonction de δ, avec $N = 10^6$.

4.3 Diffusion de rumeur en temps continu

Dans le cas du temps continu, nous supposons que parmi les n nœuds, un seul connaît initialement la rumeur et qu'un état 0 ou 1 est associée à chaque nœud. Un nœud avec l'état 1 signifie que ce nœud connaît la rumeur et un nœud avec l'état 0 signifie qu'il n'a pas connaissance de la rumeur. Pour tout $t \in \mathbb{R}^+$, nous notons $C_t^{(i)}$ l'état (0 ou 1) du nœud i à l'instant t. Á l'instant 0, tous les $C_0^{(i)}$ sont égaux à 0 sauf un qui est égal à 1 et qui correspond au nœud connaissant initialement la rumeur.

Dans le cas du temps continu, un processus de Poisson est associé à chaque nœud. Ces n processus de Poisson sont indépendants et ont le même taux $\lambda > 0$. Lorsque le processus de Poisson associé au nœud i a un saut, ce nœud choisit aléatoirement un autre nœud j, avec une distribution donnée pour interagir avec le nœud i. Cela équivaut à considérer un seul processus de Poisson avec un taux $n\lambda$ où à chaque saut deux nœuds distincts sont choisis aléatoirement avec une distribution donnée, pour interagir. Puis comme dans le cas du temps discret, les deux nœuds changent leur état avec l'état maximal de chaque nœud. De manière similaire au cas du temps discret, nous voulons évaluer le temps nécessaire pour diffuser la rumeur, c'est-à-dire le temps nécessaire pour que tous les nœuds obtiennent la valeur 1.

Nous notons $(\tau_\ell)_{\ell\in\mathbb{N}} \in \mathbb{R}^+$ les instants des sauts successifs du processus de Poisson de taux $n\lambda$, avec $\tau_0 = 0$. Donc quand le couple (i, j) est choisi à l'instant τ_ℓ, nous avons, pour $\ell \geq 1$,

$$C_t^{(i)} = C_t^{(j)} = \max\left\{C_{\tau_{\ell-1}}^{(i)}, C_{\tau_{\ell-1}}^{(j)}\right\} \text{ et } C_t^{(m)} = C_{\tau_{\ell-1}}^{(m)} \text{ pour } m \neq i, j \text{ et } t \in [\tau_\ell, \tau_{\ell+1}[.$$

Pour tout $\ell \geq 1$, nous dénotons par X_ℓ la variable aléatoire représentant ce

choix à l'instant τ_ℓ et nous supposons que ce choix est uniforme, c'est-à-dire nous supposons que, pour tout $\ell \geq 1$, nous avons

$$\mathbb{P}\{X_\ell = (i,j)\} = \frac{1}{n(n-1)} 1_{\{i \neq j\}}, \forall i, j \in [\![1, n]\!].$$

Nous considérons la variable aléatoire Θ_n définie par

$$\Theta_n = \inf \left\{ t \geq 0 \left| \sum_{i=1}^{n} C_t^{(i)} = n \right. \right\},$$

qui représente le temps nécessaire pour tous les nœuds du système pour connaître la rumeur.

Nous introduisons le processus stochastique en temps continu $Z = \{Z_t,\ t \in \mathbb{R}^+\}$ avec $[\![1, n]\!]$ comme espace d'états, défini, pour tout $t \in \mathbb{R}^+$, par

$$Z_t = \sum_{i=1}^{n} C_t^{(i)}.$$

La variable aléatoire Z_t représente le nombre de nœuds connaissant la rumeur à l'instant t. Le processus stochastique Z est alors une chaîne de Markov homogène avec une matrice des taux de transition notée B. Les valeurs non nulles de la matrice B sont données, pour $i \in [\![1, n]\!]$, par

$$\begin{cases} B_{i,i} &= -n\lambda p_i, \\ B_{i,i+1} &= n\lambda p_i, \text{ pour } i \neq n. \end{cases}$$

En effet, lorsque $Z_t = i$, le nœud suivant est activé avec l'intensité $n\lambda$. Pour que le processus Z atteigne l'état $i+1$ à partir de l'état i, soit ce nœud activé possède la rumeur (probabilité i/n) et le nœud contacté ne la possède pas (probabilité $(n-i)/(n-1)$), soit le nœud activé ne possède pas la rumeur (probabilité $(n-i)/n$) et il contacte un nœud qui la possède (probabilité $i/(n-1)$). Cela signifie que pour $i \in [\![1, n-1]\!]$, le taux de transition $B_{i,i+1}$ est donné par

$$B_{i,i+1} = n\lambda \frac{2i(n-i)}{n(n-1)} = n\lambda p_i.$$

Les états $1, \ldots, n-1$ de Z sont transitoires et l'état n est absorbant. La variable aléatoire Θ_n peut être écrite comme

$$\Theta_n = \inf\{t \geq 0 \mid Z_t = n\}.$$

Il est bien connu, voir par exemple [72], que la distribution de Θ_n est donnée, pour tout $t \geq 0$, par

$$\mathbb{P}\{\Theta_n > t\} = \alpha e^{Rt} \mathbb{1}, \tag{4.8}$$

où α est le vecteur ligne contenant les probabilités initiales des états $1, \ldots, n-1$, c'est-à-dire $\alpha_i = \mathbb{P}\{Z_0 = i\} = 1_{\{i=1\}}$, R est la sous-matrice obtenue à partir de B en supprimant la ligne et la colonne correspondant à l'état absorbant n et $\mathbb{1}$ est le

vecteur colonne de dimension $n-1$ dont toutes les composantes sont égales à 1. Pour tout $i \in [\![1, n-1]\!]$, nous notons U_i le temps de séjour du processus Z dans l'état i, qui est le temps pendant lequel il y a exactement i nœuds connaissant la rumeur. Les variables aléatoires U_i sont indépendantes et suivent la loi exponentielle de paramètre $\mu_i = n\lambda p_i$. De plus nous avons

$$\Theta_n = \sum_{i=1}^{n-1} U_i.$$

4.3.1 Espérance et variance de Θ_n

L'espérance et la variance de Θ_n ont été obtenues par S. Molchanov et J. M. Whitmeyer [61] dans le cas du modèle "push". Nous étendons ces résultats au modèle "push-pull" dans les deux lemmes suivants.

Lemme 4.3.1 *Pour tout $n \geq 2$, nous avons*

$$\mathbb{E}(\Theta_n) = \frac{(n-1)H_{n-1}}{n\lambda} \text{ et } \mathbb{E}(\Theta_n) \underset{n\longrightarrow\infty}{\sim} \frac{\ln(n)}{\lambda}.$$

Preuve. Nous avons

$$\mathbb{E}(\Theta_n) = \sum_{i=1}^{n-1} \mathbb{E}(U_i) = \frac{1}{n\lambda} \sum_{i=1}^{n-1} \frac{1}{p_i} = \frac{1}{n\lambda}\mathbb{E}(T_n) = \frac{(n-1)H_{n-1}}{n\lambda}.$$

Le reste de la preuve est évident car $H_{n-1} \underset{n\longrightarrow\infty}{\sim} \ln(n)$. ■

Lemme 4.3.2 *Pour tout $n \geq 2$, nous avons*

$$\mathbb{V}(\Theta_n) = \frac{(n-1)^2}{2n^2\lambda^2}\left(\sum_{i=1}^{n-1}\frac{1}{i^2} + \frac{2H_{n-1}}{n}\right) \leq \frac{1}{\lambda^2}\left(\frac{\pi^2}{12} + \frac{H_{n-1}}{n}\right)$$

et

$$lim_{n\longrightarrow\infty}\mathbb{V}(\Theta_n) = \frac{\pi^2}{12\lambda^2}.$$

Preuve. Les variables aléatoires U_ℓ étant indépendantes, nous avons

$$\begin{aligned}
\mathbb{V}(\Theta_n) &= \sum_{i=1}^{n-1} \mathbb{V}(U_i) = \frac{1}{n^2\lambda^2} \sum_{i=1}^{n-1} \frac{1}{p_i^2} = \frac{(n-1)^2}{4\lambda^2} \sum_{i=1}^{n-1} \frac{1}{i^2(n-i)^2} \\
&= \frac{(n-1)^2}{4n^2\lambda^2} \sum_{i=1}^{n-1} \left(\frac{1}{i} + \frac{1}{n-i} \right)^2 \\
&= \frac{(n-1)^2}{4n^2\lambda^2} \left(\sum_{i=1}^{n-1} \frac{1}{i^2} + \sum_{i=1}^{n-1} \frac{1}{(n-i)^2} + 2 \sum_{i=1}^{n-1} \frac{1}{i(n-i)} \right) \\
&= \frac{(n-1)^2}{4n^2\lambda^2} \left(2 \sum_{i=1}^{n-1} \frac{1}{i^2} + \frac{2}{n} \sum_{i=1}^{n-1} \left(\frac{1}{i} + \frac{1}{n-i} \right) \right) \\
&= \frac{(n-1)^2}{4n^2\lambda^2} \left(2 \sum_{i=1}^{n-1} \frac{1}{i^2} + \frac{4H_{n-1}}{n} \right) \\
&\leq \frac{1}{\lambda^2} \left(\frac{\pi^2}{12} + \frac{H_{n-1}}{n} \right).
\end{aligned}$$

Le reste de la preuve est évident car $H_{n-1} \underset{n\longrightarrow\infty}{\sim} \ln(n)$. ■

Notons que la différence entre le modèle push et le modèle push-pull est simplement due à un facteur de 2 dans la probabilité de transition, ce qui donne un facteur de 2 pour l'espérance et de 4 pour la variance.

4.3.2 Expression explicite de la distribution de Θ_n

La distribution de Θ_n, pour $n \geq 2$, qui est donnée par la relation (4.8) peut aisément être calculée de la manière suivante. Nous utilisons pour cela la technique d'uniformisation, voir par exemple [72]. Nous introduisons la chaîne de Markov uniformisée associée à la chaîne de Markov Z qui est caractérisée par son taux d'uniformisation ν et par sa matrice des probabilités de transition G. Le taux d'uniformisation ν doit vérifier $\nu \geq \max_{1\leq i\leq n}(-B_{i,i})$ et la matrice G est liée à B par

$$G = I_n + B/\nu,$$

où I_n est la matrice identité d'ordre n. Nous notons N_t le nombre de transitions qui ont eu lieu durant l'intervalle $[0, t]$. Le processus N_t est un processus de Poisson de taux ν et comme $B = -\nu(I_n - G)$, nous avons $R = -\nu(I_{n-1} - P)$, où P est la sous-matrice obtenue à partir de G en supprimant les lignes et les colonnes correspondant à l'état absorbant n. La relation (4.8) peut donc être réécrite de la façon suivante

$$\mathbb{P}\{\Theta_n > t\} = \alpha e^{Rt}\mathbb{1} = \sum_{k=0}^{\infty} e^{-\nu t} \frac{(\nu t)^k}{k!} \alpha P^k \mathbb{1}.$$

On vérifie facilement que

$$\max_{i\in\{1,\ldots,n\}} (-R_{i,i}) = \max_{i\in\{1,\ldots,n\}} (n\lambda p_i) \leq n\lambda.$$

En prenant $\nu = n\lambda$, nous pouvons faire un lien avec les résultats obtenus dans le cas du temps discret. En effet, nous obtenons, à partir de la relation (4.2), $P = Q$ et par conséquent, en utilisant (4.1), nous obtenons

$$\mathbb{P}\{\Theta_n > t\} = \sum_{k=0}^{\infty} e^{-n\lambda t}\frac{(n\lambda t)^k}{k!}\mathbb{P}\{T_n > k\} = \sum_{k=0}^{\infty} e^{-n\lambda t}\frac{(n\lambda t)^k}{k!}\alpha Q^k \mathbb{1}. \tag{4.9}$$

En utilisant cette expression et les résultats obtenus en temps discret nous obtenons une expression explicite de la distribution de Θ_n.

Théorème 4.3.3 *Pour tout $n \geq 1$, $t \geq 0$, nous avons*

$$\mathbb{P}\{\Theta_n > t\} = \sum_{j=1}^{\lfloor n/2 \rfloor} \left(c_{n-1,j} + n\lambda t d_{n-1,j}\right) e^{-n\lambda p_j t},$$

où les coefficients $c_{n-1,j}$ et $d_{n-1,j}$ sont donnés par les relations de récurrence (4.6).

Preuve. A partir du théorème 4.2.3, nous avons pour tout $n \geq 1$ et $k \geq 0$,

$$\mathbb{P}\{T_n > k\} = \sum_{j=1}^{\lfloor n/2 \rfloor} \left(c_{n-1,j}(1-p_j) + k d_{n-1,j}\right)(1-p_j)^{k-1},$$

où les coefficients $c_{n-1,j}$ et $d_{n-1,j}$ sont donnés par les relations (4.6). En utilisant maintenant la relation (4.9), nous obtenons

$$\begin{aligned}\mathbb{P}\{\Theta_n > t\} &= \sum_{k=0}^{\infty} e^{-n\lambda t}\frac{(n\lambda t)^k}{k!}\left(\sum_{j=1}^{\lfloor n/2 \rfloor} c_{n-1,j}(1-p_j)^k + \sum_{j=1}^{\lfloor n/2 \rfloor} k d_{n-1,j}(1-p_j)^{k-1}\right) \\ &= \sum_{j=1}^{\lfloor n/2 \rfloor} c_{n-1,j} e^{-n\lambda p_j t} + n\lambda t \sum_{j=1}^{\lfloor n/2 \rfloor} d_{n-1,j} e^{-n\lambda p_j t},\end{aligned}$$

ce qu'il fallait démontrer. ■

4.3.3 Bornes et queue de la distribution de Θ_n

Dans cette section, nous obtenons une borne très simple de la distribution de Θ_n et nous montrons que cette borne est aussi un équivalent de la queue de sa distribution. Le calcul de cette borne et de cet équivalent de la quantité $\mathbb{P}\{\Theta_n > t\}$ s'appuie sur le théorème 4.2.4 correspondant au cas du temps discret.

Théorème 4.3.4 *Pour tout $n \geq 3$ et $t \geq 0$ nous avons*

$$\mathbb{P}\{\Theta_n > t\} \leq \left[2(n-2)^2\lambda t + \frac{n}{n-2}\right] e^{-2\lambda t},$$

$$\mathbb{P}\{\Theta_n > t\} \underset{t\longrightarrow\infty}{\sim} \left[2(n-2)^2\lambda t + \frac{n}{n-2}\right] e^{-2\lambda t}.$$

Notons que pour $n = 2$, nous avons $\Theta_2 = U_1$ qui suit la loi exponentielle de paramètre $\mu_1 = 2\lambda$ et par conséquent $\mathbb{P}\{\Theta_2 > t\} = e^{-2\lambda t}$.

Preuve. A partir du théorème 4.2.4, nous avons pour $n \geq 2$ et $k \geq 1$,

$$\mathbb{P}\{T_n > k\} \leq \left(1 + \frac{2k(n-2)^2}{n}\right)\left(1 - \frac{2}{n}\right)^{k-1}.$$

Comme $\mathbb{P}\{T_n > 0\} = 1$, cela mène à

$$\begin{aligned}
\mathbb{P}\{\Theta_n > t\} &= \sum_{k=0}^{\infty} e^{-n\lambda t}\frac{(n\lambda t)^k}{k!}\mathbb{P}\{T_n > k\} \\
&\leq e^{-n\lambda t} + \sum_{k=1}^{\infty} e^{-n\lambda t}\frac{(n\lambda t)^k}{k!}\left(1 + \frac{2k(n-2)^2}{n}\right)\left(1 - \frac{2}{n}\right)^{k-1} \\
&= e^{-n\lambda t} + \sum_{k=1}^{\infty} e^{-n\lambda t}\frac{(n\lambda t)^k}{k!}\left(1 - \frac{2}{n}\right)^{k-1} \\
&\qquad + 2(n-2)^2\lambda t\sum_{k=1}^{\infty} e^{-n\lambda t}\frac{((n-2)\lambda t)^{k-1}}{(k-1)!} \\
&= e^{-n\lambda t} + \frac{ne^{-n\lambda t}\left(e^{(n-2)\lambda t} - 1\right)}{n-2} + 2(n-2)^2\lambda t e^{-n\lambda t}e^{(n-2)\lambda t} \\
&= \left[2(n-2)^2\lambda t + \frac{n}{n-2}\right]e^{-2\lambda t} - \frac{2}{n-2}e^{-n\lambda t} \\
&\leq \left[2(n-2)^2\lambda t + \frac{n}{n-2}\right]e^{-2\lambda t},
\end{aligned}$$

ce qui termine la première partie de la démonstration.

D'un coté, du fait que $p_1 < p_j$ pour $j \in \{2, \ldots, \lfloor n/2 \rfloor\}$, nous avons, à partir du théorème 4.2.3,

$$\mathbb{P}\{T_n > k\} \underset{k\longrightarrow\infty}{\sim} d_{n-1,1}k\left(1 - \frac{2}{n}\right)^{k-1}.$$

D'un autre coté, à partir du théorème 4.2.4, nous avons

$$\mathbb{P}\{T_n > k\} \underset{k\longrightarrow\infty}{\sim} \left(1 + \frac{2k(n-2)^2}{n}\right)\left(1 - \frac{2}{n}\right)^{k-1}.$$

Ces deux résultats impliquent que

$$d_{n-1,1} = \frac{2(n-2)^2}{n}.$$

De manière similaire, à partir du théorème 4.3.3, nous obtenons

$$\mathbb{P}\{\Theta_n > t\} \underset{t\longrightarrow\infty}{\sim} d_{n-1,1}n\lambda t e^{-n\lambda p_1 t} = 2(n-2)^2\lambda t e^{-2\lambda t},$$

et aussi

$$2(n-2)^2\lambda t e^{-2\lambda t} \underset{t\longrightarrow\infty}{\sim} \left[2(n-2)^2\lambda t + \frac{n}{n-2}\right]e^{-2\lambda t},$$

ce qui termine la preuve. ■

Nous donnons dans ce qui suit deux bornes différentes pour la quantité $\mathbb{P}\{\Theta_n > c\mathbb{E}(\Theta_n)\}$, avec $c \geq 1$. Ces bornes seront comparées et utilisées pour obtenir le comportement de cette quantité quand le nombre de nœuds n tend vers l'infini.

En rappelant que $\mathbb{E}(\Theta_n) = (n-1)H_{n-1}/(n\lambda)$, une première borne est obtenue par une application immédiate du théorème 5.1 de [44], qui mène, pour tout $n \geq 3$ et pour tout nombre réel $c \geq 1$, à

$$\mathbb{P}\{\Theta_n > c\mathbb{E}(\Theta_n)\} \leq \frac{1}{c}\exp\left(-\frac{2(n-1)H_{n-1}(c-1-\ln(c))}{n}\right). \tag{4.10}$$

Notons que le terme de droite est égal à 1 quand $c = 1$.

En appliquant le théorème 4.3.4 au point $c\mathbb{E}(\Theta_n)$, nous obtenons la seconde borne suivante.

$$\begin{aligned}\mathbb{P}\{\Theta_n > c\mathbb{E}(\Theta_n)\} &\leq \left[2(n-2)^2\lambda c\mathbb{E}(\Theta_n) + \frac{n}{n-2}\right] e^{-2\lambda c\mathbb{E}(\Theta_n)} \\ &= \left[\frac{2c(n-2)^2(n-1)H_{n-1}}{n} + \frac{n}{n-2}\right]\exp\left(-\frac{2c(n-1)H_{n-1}}{n}\right).\end{aligned}$$

À partir de maintenant, nous notons cette borne $\varphi(c,n)$ et, de la même manière, nous notons $\psi(c,n)$ la borne de $\mathbb{P}\{\Theta_n > c\mathbb{E}(\Theta_n)\}$ obtenue dans (4.10). Nous avons alors, pour $n \geq 3$ et $c \geq 1$,

$$\begin{aligned}\varphi(c,n) &= \left[\frac{2c(n-2)^2(n-1)H_{n-1}}{n} + \frac{n}{n-2}\right]\exp\left(-\frac{2c(n-1)H_{n-1}}{n}\right), \\ \psi(c,n) &= \frac{1}{c}\exp\left(-\frac{2(n-1)H_{n-1}(c-1-\ln(c))}{n}\right).\end{aligned}$$

Ces deux bornes sont comparées dans le théorème suivant.

Théorème 4.3.5 *Pour tout $n \geq 5$, il existe un unique $c^* \geq 1$ tel que $\varphi(c^*,n) = \psi(c^*,n)$ et nous avons*

$$\begin{cases} \varphi(c,n) > \psi(c,n) & \text{for all } 1 \leq c < c^* \\ \varphi(c,n) < \psi(c,n) & \text{for all } c > c^*. \end{cases} \tag{4.11}$$

De plus,

$$\lim_{c\longrightarrow\infty}\frac{\varphi(c,n)}{\psi(c,n)} = 0.$$

Preuve. Nous introduisons les quantités

$$A_n = \frac{(n-1)H_{n-1}}{n}, B_n = 2(n-2)^2A_n \text{ et } C_n = \frac{n}{n-2}.$$

Nous obtenons

$$\frac{\varphi(c,n)}{\psi(c,n)} = \left(B_nc^2 + C_nc\right)e^{-2A_n(1+\ln(c))} = \left(B_nc^{2-2A_n} + C_nc^{1-2A_n}\right)e^{-2A_n}.$$

On vérifie aisément que la suite A_n est strictement croissante et que $A_3 = 1$. Il s'ensuit que pour $n \geq 5$, nous avons $A_n > 1$ et aussi

$$1 - 2A_n < 2 - 2A_n < 0,$$

ceci implique que pour tout $n \geq 5$, la fonction $\varphi(c,n)/\psi(c,n)$ est strictement décroissante en fonction de c sur $[1, +\infty[$ et que

$$\lim_{c \longrightarrow \infty} \frac{\varphi(c,n)}{\psi(c,n)} = 0.$$

Considérons maintenant les suites x_n et y_n définies pour $n \geq 5$, par

$$x_n = \frac{\varphi(1,n)}{\psi(1,n)} = (B_n + C_n)\, e^{-2A_n} \text{ et } y_n = \frac{2e^{-2}(n-2)^2 A_n}{(n-1)^2}.$$

La suite A_n étant croissante, on vérifie aisément que la suite y_n est aussi croissante. De plus, nous avons

$$x_n \geq B_n e^{-2(1+\ln(n-1))} = \frac{e^{-2}B_n}{(n-1)^2} = \frac{2e^{-2}(n-2)^2 A_n}{(n-1)^2} = y_n.$$

Un simple calcul montre que nous avons $y_{34} > 1$. La suite y_n étant croissante, nous obtenons $y_n > 1$ pour tout $n \geq 34$. Il s'ensuit que nous avons $x_n > 1$ pour tout $n \geq 34$. Un calcul numérique montre que $x_n > 1$ pour $n \in \{5, \ldots, 33\}$ ce qui veut dire que pour tout $n \geq 5$, nous avons $x_n = \varphi(1,n)/\psi(1,n) > 1$. La fonction $\varphi(c,n)/\psi(c,n)$ étant strictement décroissante en fonction de c sur $[1, +\infty[$, nous en déduisons qu'il existe une unique solution, que nous notons c^*, à l'équation $\varphi(c,n)/\psi(c,n) = 1$ et (4.11) en découle. ■

Ce théorème montre que notre borne $\varphi(c,n)$ est bien meilleure que celle obtenue en utilisant le résultat de [44], qui avait été noté $\psi(c,n)$, pour $c > c^*$, non seulement parce que le ratio $\varphi(c,n)/\psi(c,n)$ décroît avec c et tend vers 0 quand c tend vers l'infini, mais aussi parce que, pour tout valeur de n, la valeur de c^* est très proche de 1 comme le montre le tableau 4.2. En outre, à partir du théorème 4.3.4, notre borne est optimale dans la mesure où

$$\mathbb{P}\{\Theta_n > c\mathbb{E}(\Theta_n)\} \underset{c \longrightarrow \infty}{\sim} \varphi(c,n).$$

Tableau 4.2 : Valeur de c^* pour différentes tailles du système n.

n	10	10^2	10^3	10^4	10^5	10^6	10^7	10^8	10^9
c^*	1.253	1.163	1.128	1.109	1.095	1.085	1.078	1.071	1.066

Le tableau 4.3 et la figure 4.7 illustrent, pour un système constitué de $n = 1000$ nœuds, le comportement des bornes $\varphi(c, 1000)$ et $\psi(c, 1000)$, en fonction de c, comparé à la valeur exacte de la distribution de Θ_{1000} aux points $c\mathbb{E}(\Theta_{1000})$, calculée en utilisant le théorème 4.3.3. Le tableau 4.3 illustre clairement le résultat

Tableau 4.3 : Valeurs de $\mathbb{P}\{\Theta_{1000} > c\mathbb{E}(\Theta_{1000})\}$, $\varphi(c, 1000)$ et $\psi(c, 1000)$ pour différentes valeurs de c.

c	1	1.2	1.4	1.6	1.8	2
$\mathbb{P}\{\Theta_{1000} > c\mathbb{E}(\Theta_{1000})\}$	0.446	0.063	0.005	3.9×10^{-4}	2.6×10^{-5}	1.6×10^{-6}
$\varphi(c, 1000)$	≥ 1	0.288	0.017	9.7×10^{-4}	5.5×10^{-5}	3×10^{-6}
$\psi(c, 1000)$	1.0	0.634	0.276	0.089	0.023	0.005

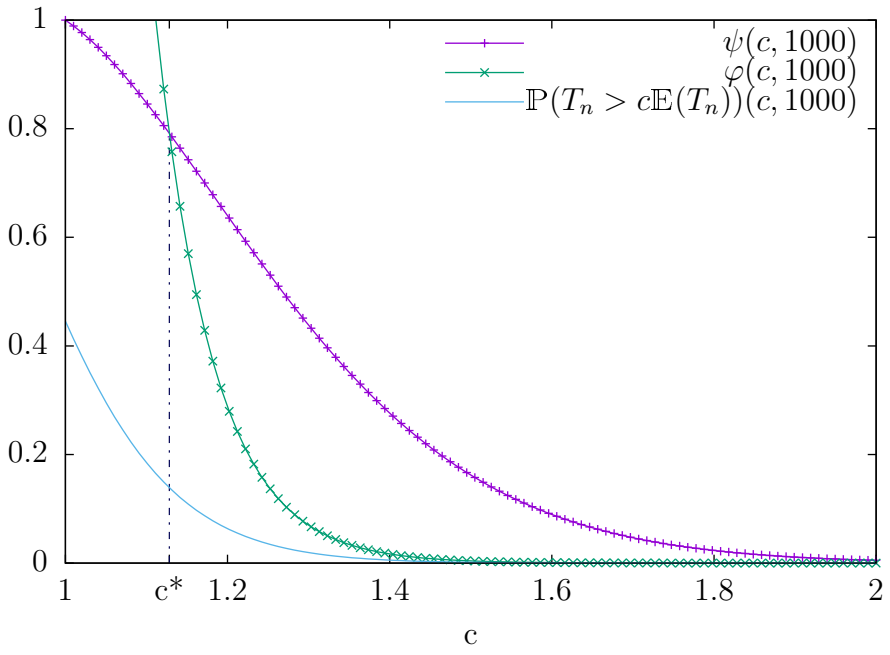


FIGURE 4.7 : Bornes $\psi(c, 1000)$, $\varphi(c, 1000)$ et valeur réelle de $\mathbb{P}\{\Theta_{1000} > c\mathbb{E}(\Theta_{1000}\}$ en fonction de c. Le point où les bornes sont égales est $c^* = 1, 12819634$.

du théorème 4.3.4. En effet, les valeurs de notre borne $\varphi(c, 1000)$ sont très proches des valeurs réelles de la distribution, tandis que les valeurs de $\psi(c, 1000)$ ont tendance à s'écarter de la valeur réelle même pour des valeurs de c peu élevées. Notons que pour $c = 1$ les deux bornes sont sans intérêt et la valeur réelle de $\mathbb{P}\{\Theta_{1000} > \mathbb{E}(\Theta_{1000})\}$ est très proche de la limite obtenue dans le théorème 4.3.8 de la section suivante. La figure 4.7 montre le grand écart entre les bornes $\varphi(c, 1000)$ et $\psi(c, 1000)$ quand c est plus grand que c^*. Cet écart grandit en valeur relative en fonction de c comme on peut le voir dans le tableau 4.3. De plus ce large écart grandit en fonction de n car la valeur de c^* semble décroître vers 1 quand n croît, comme on peut le voir dans la tableau 4.2.

4.3.4 Analyse asymptotique de la distribution de Θ_n

Nous analysons dans cette section le comportement de la distribution de Θ_n au point $c\mathbb{E}(\Theta_n)$ quand le nombre n de nœuds dans le système tend vers l'infini, en fonction de c.

Nous prouvons dans le théorème suivant que les bornes $\varphi(c, n)$ et $\psi(c, n)$, ob-

tenues à partir du théorème 4.3.4 et de la relation (4.10) respectivement avec $t = c\mathbb{E}(T_n)$, tendent toutes les deux vers 0 quand n tend vers l'infini.

Théorème 4.3.6 *Pour tout nombre réel $c \neq 1$, nous avons*

$$\lim_{n\longrightarrow\infty} \varphi(c,n) = 0 \text{ et } \lim_{n\longrightarrow\infty} \psi(c,n) = 0.$$

Preuve. On vérifie facilement que

$$\varphi(c,n) \underset{n\longrightarrow\infty}{\sim} \frac{2cn^2\ln(n)}{n^{2c}}$$

qui tend vers 0 quand n tend vers l'infini. Concernant $\psi(c,n)$ nous avons

$$\psi(c,n) \underset{n\longrightarrow\infty}{\sim} \frac{1}{c}e^{-\ln(n)(c-1-\ln(c))}.$$

Pour $c > 1$ nous avons $c - 1 - \ln(c) > 0$ ce qui implique que $\psi(c,n)$ tend vers 0 quand n tend vers l'infini. ■

Théorème 4.3.7 *Pour tout réel $c \geq 0$, nous avons*

$$\lim_{n\to\infty} \mathbb{P}\{\Theta_n > c\mathbb{E}(\Theta_n)\} = \begin{cases} 0 & \text{if } c > 1 \\ 1 & \text{if } c < 1. \end{cases}$$

Preuve. A partir du théorème 4.3.6, les deux bornes $\varphi(c,n)$ et $\psi(c,n)$ de $\mathbb{P}\{\Theta_n > c\mathbb{E}(\Theta_n)\}$ tendent vers 0 quand n tend vers l'infini, pour $c > 1$. Aussi en utilisant $\varphi(c,n)$ ou $\psi(c,n)$ nous pouvons en déduire que

$$\lim_{n\longrightarrow\infty} \mathbb{P}\{\Theta_n > c\mathbb{E}(\Theta_n)\} = 0 \text{ pour tout } c > 1.$$

Dans le cas où $c < 1$, le théorème 5.1 de [44] mène à

$$\mathbb{P}\{\Theta_n > c\mathbb{E}(\Theta_n)\} \geq 1 - \exp\left(\frac{-2(n-1)H_{n-1}(c-1-\ln(c))}{n}\right).$$

Comme $c-1-\ln(c) > 0$ pour tout $c \in [0,1[$, le terme de droite de cette inégalité tend vers 1 quand $n \longrightarrow \infty$. Par conséquent, $\lim_{n\longrightarrow\infty} \mathbb{P}\{\Theta_n > c\mathbb{E}(\Theta_n)\} = 1$ quand $c < 1$. ■

Le théorème suivant concerne le cas où $c = 1$. Notons que le résultat est identique à celui du théorème 4.2.10 dans le cas du temps discret.

Théorème 4.3.8

$$\lim_{n\longrightarrow\infty} \mathbb{P}\{\Theta_n > \mathbb{E}(\Theta_n)\} = 1 - 2e^{-\gamma}K_1\left(2e^{-\gamma}\right) \approx 0.448429663727,$$

où γ est la constante d'Euler donnée par $\gamma = \lim_{n\longrightarrow\infty}(H_n - \ln(n)) \approx 0.5772156649$ et K_1 est la fonction de Bessel modifiée de deuxième espèce et d'ordre 1, donnée pour $z > 0$, par

$$K_1(z) = \frac{z}{4}\int_0^{+\infty} t^{-2}e^{-t-z^2/4t}dt.$$

Preuve. Voir annexe section A.2 ■

Remarque. Voici quelques développements possibles pour la suite.

1. Dans le cas continu, nous avons supposé que le nombre initial de nœuds connaissant la rumeur est égal à 1. Le cas où ce nombre est égal à ℓ, avec $\ell \geq 2$, a été traité dans le cas discret. Cette extension au cas du temps continu est probablement simple car il suffit de redéfinir la variable aléatoire Θ_n comme $\Theta_n = U_\ell + \cdots + U_n$.

2. Au lieu de considérer le temps total nécessaire pour que tous les nœuds obtiennent la rumeur, nous pouvons considérer le temps total nécessaire pour qu'un pourcentage fixe, disons ρ, des nœuds obtiennent la rumeur. Dans ce cas, la variable aléatoire Θ_n à considérer devrait être redéfinie comme suit : $\Theta_n = U_1 + \cdots + U_{\lceil \rho n \rceil}$. Bien entendu, cette extension pourrait également être combinée avec la première ci-dessus.

3. Les instants auxquels se produisent les interactions entre les nœuds ont été modélisés par un processus de Poisson. Cela pourrait être généralisé en considérant, au lieu d'un processus de Poisson, un processus de renouvellement de type Phase qui préserve la propriété Markov et peut se rapprocher de chaque processus ponctuel.

Remerciements. Nous tenons à remercier le professeur Philippe Carmona pour ses conseils d'experts concernant la preuve du théorème 4.2.10.

4.4 Conclusion

Dans ce chapitre, nous avons fourni une analyse approfondie du temps de propagation de la rumeur dans le modèle push-pull asynchrone dans le cas du temps continu en complétant et en étendant les résultats déjà obtenus dans le cas du temps discret. Une telle analyse précise est un pas en avant vers la conception de solutions plus complexes de problèmes tels que, par exemple, l'élection du leader dans les grands systèmes distribués. Notre analyse concernant la queue de la distribution du temps de propagation de la rumeur et son comportement limite lorsque le nombre de nœuds tend vers l'infini a été réalisée avec une grande précision ouvrant la voie à des applications pratiques. Le chapitre 7 en fourni un exemple.

Chapitre 5

Protocoles basés sur la moyenne

Dans ce chapitre, nous nous intéressons aux protocoles basés sur la moyenne, protocoles qui peuvent avoir trois applications : le problème de la proportion, le problème du comptage et le problème de la majorité. Ces protocoles ont déjà été présentés en sections 3.3, 3.2 et 3.4. Les protocoles basés sur la moyenne ont deux variantes, une où les états sont des réels et l'autre où les états sont des entiers. Il y a des similitudes mais aussi de grandes différences dans le traitement des démonstrations, c'est pourquoi nous étudierons ces sujets dans deux sections différentes.

5.1 Protocoles de moyenne avec des réels

Ce chapitre aborde les problèmes de la proportion, du comptage et de la majorité avec des protocoles dont les états sont des nombres réels. Son originalité est d'utiliser, en plus de la norme euclidienne, la norme 4, ce qui complexifie les calculs, mais mène à des résultats plus fins.

Nous avons un nombre infini d'états, donc nous ne parlerons pas de Protocoles de Population. Ceci étant dit, rien ne nous empêche d'utiliser un formalisme identique, et c'est ce que nous ferons. Il faut noter que Sauerwald et Sun en 2012 [71] ont montré un lien entre les performances obtenues avec des réels et celles obtenues avec des entiers, ce qui veut dire que les résultats obtenus dans ce chapitre peuvent avoir une application dans le domaine des protocoles de population au sens strict, c'est-à-dire avec un nombre fini d'états.

Ce chapitre reprend les résultats de deux de nos publications [54, 58].

5.1.1 Les protocoles de proportion, de comptage et de majorité

Les problèmes de la proportion, du comptage et de la majorité ont été définis respectivement en sections 3.3 et 3.2. Nous rappelons que n est la taille du système, n_A (respectivement n_B) est le nombre de nœuds initialement à A (respectivement B), $\gamma_A = n_A/n$ (respectivement $\gamma_B = n_B/n$) est la proportion de nœuds initialement à A (respectivement B).

En ce qui concerne les solutions, nous utilisons le même formalisme que pour les protocoles de population (voir section 2.2). Le protocole de proportion est dé-

fini par le sextuplet $(\Sigma, \Xi, Q, \iota, \omega_A, f)$, le protocole de comptage est défini par le sextuplet $(\Sigma, \Xi', Q, \iota, \omega'_A, f)$ et le protocole de majorité est défini par le sextuplet $(\Sigma, \Xi'', Q, \iota, \omega''_A, f)$. Comme nous pouvons le constater, les seules différences entre les protocoles se situent au niveau de la fonction de sortie et au niveau de l'ensemble de sortie.

Initialement, tous les nœuds commencent avec le symbole A ou B qui fournit leur état initial, donc nous avons $\Sigma = \{A, B\}$. Soit m un nombre réel strictement positif. La fonction d'entrée ι attribue aux agents la valeur m ou $-m$ selon que leur symbole est A ou B. Nous avons donc $\iota(A) = m$ et $\iota(B) = -m$. La moyenne étant effectuée entre les agents, l'ensemble des états est $Q = [-m, m]$. Notons que cet ensemble est infini. La fonction de transition f est ainsi définie

$$\begin{aligned} f : [-m, m] \times [-m, m] &\longrightarrow [-m, m] \times [-m, m] \\ (x, y) &\longmapsto \left(\frac{x+y}{2}, \frac{x+y}{2} \right). \end{aligned}$$

Pour le protocole de proportion, on cherche à connaître la proportion initiale de A dans le système. L'ensemble de sortie est $\Xi = [0, 1]$, la fonction de sortie est définie par

$$\begin{aligned} \omega_A : [-m, m] &\longrightarrow [0, 1] \\ x &\longmapsto \frac{x+m}{2m}. \end{aligned}$$

Pour le protocole de comptage, on cherche à connaître le nombre de nœuds initialisés à A. L'ensemble de sortie est naturellement l'ensemble des entiers de 0 à n, c'est-à-dire $\Xi' = [\![0, n]\!]$, la fonction de sortie est définie par

$$\begin{aligned} \omega'_A : [-m, m] &\longrightarrow [\![0, n]\!] \\ x &\longmapsto \left\lfloor \frac{n(x+m)}{2m} + \frac{1}{2} \right\rfloor. \end{aligned}$$

Pour le protocole de majorité, on cherche à savoir si le nombre de nœuds initialisés à A est supérieur ou non au nombre de nœuds initialisé à B. L'ensemble de sortie est la valeur booléenne de la réponse 1 pour oui et 0 pour non, c'est-à-dire $\Xi'' = \{0, 1\}$. La fonction de sortie est définie par

$$\begin{aligned} \omega''_A : [-m, m] &\longrightarrow \{0, 1\} \\ x &\longmapsto 1_{\{x \geq 0\}}. \end{aligned}$$

Nous utilisons les même notations que celles indiquées dans la section 2.2. À chaque instant $t \geq 0$, la configuration t du protocole est notée $C_t = (C_t^{(1)}, \ldots, C_t^{(n)})$, où $C_t^{(i)}$ est l'état du nœud i à l'instant t. $C = \{C_t,\ t \geq 0\}$ est un processus stochastique en temps discret sur l'espace d'états $[-m, m]^n$. Les interactions entre les nœuds sont orchestrées par un ordonnanceur aléatoire : à chaque instant discret $t \geq 0$, deux indices quelconques i et j sont choisis au hasard pour interagir avec une probabilité $p_{i,j}(t)$. Une fois choisi, le couple de nœuds (i, j) interagit et les deux nœuds mettent

à jour leur état respectif $C_t^{(i)}$ et $C_t^{(j)}$ en appliquant la fonction de transition f, définie précédemment, menant à l'état C_{t+1}. Nous avons donc

$$\left(C_{t+1}^{(i)}, C_{t+1}^{(j)}\right) = f\left(C_t^{(i)}, C_t^{(j)}\right) = \left(\frac{C_t^{(i)} + C_t^{(j)}}{2}, \frac{C_t^{(i)} + C_t^{(j)}}{2}\right) \tag{5.1}$$
$$\text{et } C_{t+1}^{(r)} = C_t^{(r)} \text{ pour } r \neq i, j.$$

Nous notons X_t la variable aléatoire correspondant au choix du couple d'agents devant interagir et nous supposons que X_t est uniformément distribuée, c'est-à-dire

$$\mathbb{P}\left\{X_t = (i,j)\right\} = p_{i,j}(t) = \frac{1_{\{i \neq j\}}}{n(n-1)}.$$

À chaque instant t et à la demande de l'application, tout nœud i du système peut fournir son estimation de $\gamma_A = n_A/n$ en renvoyant $\omega_A\left(C_t^{(i)}\right)$, c'est-à-dire

$$\omega_A\left(C_t^{(i)}\right) = \frac{C_t^{(i)} + m}{2m}.$$

À chaque instant t, et à la demande de l'application, tout nœud i du système peut fournir son estimation de n_A en renvoyant $\omega'_A\left(C_t^{(i)}\right)$, c'est-à-dire

$$\omega'_A\left(C_t^{(i)}\right) = \left\lfloor \frac{C_t^{(i)} n}{m} + \frac{1}{2} \right\rfloor.$$

À chaque instant t, et à la demande de l'application, tout nœud i du système peut fournir son estimation de la réponse à la question : "A est-il majoritaire ?" en renvoyant $\omega''_A\left(C_t^{(i)}\right)$, c'est-à-dire

$$\omega''_A\left(C_t^{(i)}\right) = 1_{\left\{C_t^{(i)} > 0\right\}}.$$

Nous montrons avec le corollaire 5.1.8 (dans la section 5.1.3) qu'après un certain temps explicitement donné, avec une probabilité élevée et avec une précision élevée, $\omega_A\left(C_t^{(i)}\right) \simeq \gamma_A = n_A/n$ pour tout nœud i dans le système.

Nous montrons dans la suite (voir le corollaire 5.1.9) qu'après un certain temps explicitement donné et avec une probabilité élevée, $\omega'_A\left(C_t^{(i)}\right) = n_A$ pour tout nœud i dans le système.

Nous montrons aussi (voir le corollaire 5.1.10) qu'après un certain temps explicitement donné et avec une probabilité élevée, $\omega''_A\left(C_t^{(i)}\right) = 1_{\{n_A > n_B\}}$ pour tout nœud i.

Notons que le nœud i n'a pas besoin de connaître la taille n du système pour calculer la proportion de nœuds qui ont commencé avec le symbole A, ni pour savoir si les A sont majoritaires.

5.1.2 Résultats préliminaires

Le lemme suivant indique que la somme des entrées du vecteur C_t est constante.

Lemme 5.1.1 *Pour tout $t \geq 0$, $\sum_{i=1}^{n} C_t^{(i)} = \sum_{i=1}^{n} C_0^{(i)}$.*

On note ℓ la valeur moyenne de la somme des composantes de C_t et L le vecteur ligne de $\mathbb{R}^n$ dont toutes les composantes sont égales à ℓ, c'est-à-dire

$$\ell = \frac{1}{n} \sum_{i=1}^{n} C_t^{(i)} \text{ et } L = (\ell, \ldots, \ell).$$

Pour tout $d \in \mathbb{N} \setminus \{0\}$ et $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, nous utiliserons la norme d et la norme ∞ de x notées $\|x\|_d$ et $\|x\|_\infty$, et définies par

$$\|x\|_d = \left(\sum_{i=1}^{n} |x_i|^d \right)^{1/d} \text{ et } \|x\|_\infty = \max_{i=1,\ldots,n} |x_i|.$$

Il est bien connu que ces normes satisfont

$$\|x\|_\infty \leq \|x\|_d \leq n^{1/d} \|x\|_\infty.$$

Cela montre notamment que

$$\lim_{d \longrightarrow \infty} \|x\|_d = \|x\|_\infty. \tag{5.2}$$

Pour simplifier, nous définissons, pour cette section, la notation

$$y_t^{(i)} = C_t^{(i)} - \ell \text{ et } Y_t = (y_t^{(1)}, \ldots, y_t^{(n)}),$$

soit

$$Y_t = C_t - L.$$

Le théorème suivant indique que, pour tout d, la suite des normes d du vecteur Y_t est décroissante.

Théorème 5.1.2 *Pour tout $d \in \mathbb{N}^* \cup \{+\infty\}$, la suite $(\|Y_t\|_d)_{t \geq 0}$ est décroissante.*

Preuve. Supposons d'abord que d est fini avec $d \geq 1$. A partir de la relation (5.1), on a, pour tout $t \geq 0$,

$$\|Y_{t+1}\|_d^d = \|Y_t\|_d^d - \sum_{i,j=1}^{n} \left(\left| y_t^{(i)} \right|^d + \left| y_t^{(j)} \right|^d - 2 \left| \frac{y_t^{(i)} + y_t^{(j)}}{2} \right|^d \right) 1_{\{X_t = (i,j)\}}. \tag{5.3}$$

La fonction réelle g définie par $g(x) = x^d$ est une fonction convexe sur $[0, \infty[$, donc pour chaque $a, b \geq 0$, nous avons

$$a^d + b^d \geq 2 \left(\frac{a+b}{2} \right)^d.$$

En prenant $a = \left|y_t^{(i)}\right|$, $b = \left|y_t^{(j)}\right|$ et en utilisant le fait que $|a| + |b| \geq |a + b|$, nous obtenons

$$\left|y_t^{(i)}\right|^d + \left|y_t^{(j)}\right|^d \geq 2\left(\frac{\left|y_t^{(i)}\right| + \left|y_t^{(j)}\right|}{2}\right)^d \geq 2\left|\frac{y_t^{(i)} + y_t^{(j)}}{2}\right|^d,$$

ce qui signifie que la double somme dans (5.3) est positive. Cela prouve que $\|Y_{t+1}\|_d^d \leq \|Y_t\|_d^d$ c'est-à-dire $\|Y_{t+1}\|_d \leq \|Y_t\|_d$. Si $d = \infty$, en prenant la limite de cette inégalité, on obtient en utilisant (5.2) $\|Y_{t+1}\|_\infty \leq \|Y_t\|_\infty$, ce qui termine la preuve. ■

5.1.3 Analyse

Nous étudions en premier lieu les moments de $\|Y_t\|_2$ et $\|Y_t\|_4$ ce qui nous permettra d'analyser leur distribution. Nous supposons que $n \geq 3$.

Théorème 5.1.3 *Pour tout $t \geq 0$, nous avons*

$$\mathbb{E}\left(\|Y_t\|_2^2\right) = \left(1 - \frac{1}{n-1}\right)^t \mathbb{E}\left(\|Y_0\|_2^2\right). \tag{5.4}$$

Preuve. Soit $x \in \mathbb{R}^n$. Pour tout $i, j \in [\![1, n]\!]$ avec $i \neq j$, le vecteur $y = (y_1, \ldots, y_n)$ défini par

$$y_i = y_j = \frac{x_i + x_j}{2} \text{ et } y_r = x_r \text{ pour } r \neq i, j$$

vérifie

$$\|y - L\|_2^2 = \|x - L\|_2^2 - (x_i - \ell)^2 - (x_j - \ell)^2 + 2\left(\frac{x_i + x_j}{2} - \ell\right)^2,$$

ce qui donne

$$\|y - L\|_2^2 = \|x - L\|_2^2 - \frac{(x_i - x_j)^2}{2}.$$

En appliquant ces résultats aux vecteurs aléatoires C_{t+1} and C_t cela donne, pour tout $t \geq 0$,

$$\|C_{t+1} - L\|_2^2 = \|C_t - L\|_2^2 - \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \left(C_t^{(i)} - C_t^{(j)}\right)^2 1_{\{X_t=(i,j)\}},$$

c'est-à-dire

$$\|Y_{t+1}\|_2^2 = \|Y_t\|_2^2 - \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \left(C_t^{(i)} - C_t^{(j)}\right)^2 1_{\{X_t=(i,j)\}}. \tag{5.5}$$

En prenant l'espérance et en utilisant le fait que X_t et C_t sont indépendants, nous obtenons

$$\mathbb{E}\left(\|Y_{t+1}\|_2^2\right) = \mathbb{E}\left(\|Y_t\|_2^2\right) - \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \mathbb{E}\left(\left(C_t^{(i)} - C_t^{(j)}\right)^2\right) p_{i,j}(t).$$

Comme

$$p_{i,j}(t) = \frac{1}{n(n-1)},$$

nous avons

$$\mathbb{E}\left(\|Y_{t+1}\|_2^2\right) = \mathbb{E}\left(\|Y_t\|_2^2\right) - \frac{1}{2n(n-1)} \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{E}\left(\left(C_t^{(i)} - C_t^{(j)}\right)^2\right),$$

et de plus

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n}\left(C_t^{(i)} - C_t^{(j)}\right)^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\left(C_t^{(i)} - \ell\right) - \left(C_t^{(j)} - \ell\right)\right)^2 \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}\Big[\left(C_t^{(i)} - \ell\right)^2 + \left(C_t^{(j)} - \ell\right)^2 \\
&\qquad -2\left(C_t^{(i)} - \ell\right)\left(C_t^{(j)} - \ell\right)\Big] \\
&= 2n\|Y_t\|_2^2 - 2\sum_{i=1}^{n}\left(C_t^{(i)} - \ell\right)\sum_{j=1}^{n}\left(C_t^{(j)} - \ell\right) \\
&= 2n\|Y_t\|_2^2 - 2\sum_{i=1}^{n}\left(C_t^{(i)} - \ell\right)(n\ell - n\ell) \\
&= 2n\|Y_t\|_2^2.
\end{aligned}$$

Cela mène à

$$\mathbb{E}\left(\|Y_{t+1}\|_2^2\right) = \left(1 - \frac{1}{n-1}\right)\mathbb{E}\left(\|Y_t\|_2^2\right),$$

et par conséquent

$$\mathbb{E}\left(\|Y_t\|^2\right) = \left(1 - \frac{1}{n-1}\right)^t \mathbb{E}\left(\|Y_0\|_2^2\right),$$

ce qu'il fallait démontrer. ■

Théorème 5.1.4 *Pour tout $t \geq 0$, nous avons*

$$\mathbb{E}\left(\|Y_{t+1}\|_4^4\right) = \left(1 - \frac{7}{4(n-1)}\right)\mathbb{E}\left(\|Y_t\|_4^4\right) + \frac{3}{4n(n-1)}\mathbb{E}\left(\|Y_t\|_2^4\right). \tag{5.6}$$

Preuve. En prenant les espérances dans la relation (5.3) et en utilisant le fait que X_t et Y_t sont indépendants, nous obtenons pour $d = 4$,

$$\begin{aligned}
&\mathbb{E}\left(\|Y_{t+1}\|_4^4\right) \\
&= \mathbb{E}\left(\|Y_t\|_4^4\right) - \sum_{i,j=1}^{n}\mathbb{E}\left(\left(y_t^{(i)}\right)^4 + \left(y_t^{(j)}\right)^4 - 2\left(\frac{y_t^{(i)} + y_t^{(j)}}{2}\right)^4\right) p_{i,j}(t) \\
&= \mathbb{E}\left(\|Y_t\|_4^4\right) - \frac{1}{n(n-1)}\sum_{i,j=1}^{n}\mathbb{E}\left[\left(y_t^{(i)}\right)^4 + \left(y_t^{(j)}\right)^4 - 2\left(\frac{y_t^{(i)} + y_t^{(j)}}{2}\right)^4\right]. \qquad (5.7)
\end{aligned}$$

La double somme (5.7) peut aussi s'écrire

$$\sum_{i,j=1}^{n}\left[\left(y_t^{(i)}\right)^4+\left(y_t^{(j)}\right)^4-2\left(\frac{y_t^{(i)}+y_t^{(j)}}{2}\right)^4\right]=\frac{7}{8}\sum_{i,j=1}^{n}\left(\left(y_t^{(i)}\right)^4+\left(y_t^{(j)}\right)^4\right)-\frac{1}{2}\sum_{i,j=1}^{n}\left(y_t^{(i)}\left(y_t^{(j)}\right)^3+\left(y_t^{(i)}\right)^3y_t^{(j)}\right)-\frac{3}{4}\sum_{i,j=1}^{n}\left(y_t^{(i)}\right)^2\left(y_t^{(j)}\right)^2.$$

Nous allons considérer les 3 termes séparément. Pour le premier, nous avons

$$\frac{7}{8}\sum_{i,j=1}^{n}\left(\left(y_t^{(i)}\right)^4+\left(y_t^{(j)}\right)^4\right)=\frac{7n}{4}\sum_{i=1}^{n}\left(y_t^{(i)}\right)^4=\frac{7n}{4}\|Y_t\|_4^4.$$

Pour le second terme, par définition de ℓ, nous avons

$$\sum_{i=1}^{n}y_t^{(i)}=\sum_{i=1}^{n}C_t^{(i)}-n\ell=0 \tag{5.8}$$

ce qui nous permet d'obtenir

$$\frac{1}{2}\sum_{i,j=1}^{n}\left(y_t^{(i)}\left(y_t^{(j)}\right)^3+\left(y_t^{(i)}\right)^3y_t^{(j)}\right)=\sum_{j=1}^{n}\left(y_t^{(j)}\right)^3\sum_{i=1}^{n}\left(y_t^{(i)}\right)=0.$$

Enfin, pour le troisième terme nous avons

$$\frac{3}{4}\sum_{i,j=1}^{n}\left(y_t^{(i)}\right)^2\left(y_t^{(j)}\right)^2=\frac{3}{4}\sum_{i=1}^{n}\left(y_t^{(i)}\right)^2\sum_{j=1}^{n}\left(y_t^{(j)}\right)^2=\frac{3}{4}\|Y_t\|_2^4.$$

Nous obtenons donc

$$\sum_{i,j=1}^{n}\left[\left(y_t^{(i)}\right)^4+\left(y_t^{(j)}\right)^4-2\left(\frac{y_t^{(i)}+y_t^{(j)}}{2}\right)^4\right]=\frac{7n}{4}\|Y_t\|_4^4-\frac{3}{4}\|Y_t\|_2^4.$$

La relation (5.7) devient

$$\mathbb{E}\left(\|Y_{t+1}\|_4^4\right)=\mathbb{E}\left(\|Y_t\|_4^4\right)-\frac{7}{4(n-1)}\mathbb{E}\left(\|Y_t\|_4^4\right)+\frac{3}{4n(n-1)}\mathbb{E}\left(\|Y_t\|_2^4\right),$$

c'est-à-dire

$$\mathbb{E}\left(\|Y_{t+1}\|_4^4\right)=\left(1-\frac{7}{4(n-1)}\right)\mathbb{E}\left(\|Y_t\|_4^4\right)+\frac{3}{4n(n-1)}\mathbb{E}\left(\|Y_t\|_2^4\right),$$

ce qu'il fallait démontrer. ■

Pour que ce résultat soit vraiment intéressant, nous devons évaluer le moment d'ordre 4 de la norme 2 de Y_t. C'est l'objectif du théorème suivant.

Théorème 5.1.5 *Pour tout $t \geq 0$, nous avons*

$$\mathbb{E}\left(\|Y_{t+1}\|_2^4\right) = \left(1 - \frac{4n-3}{2n(n-1)}\right)\mathbb{E}\left(\|Y_t\|_2^4\right) + \frac{1}{2(n-1)}\mathbb{E}\left(\|Y_t\|_4^4\right). \tag{5.9}$$

Preuve. En appliquant la relation (5.3) avec $d = 2$ nous avons

$$\begin{aligned}\|Y_{t+1}\|_2^2 &= \|Y_t\|_2^2 - \sum_{i,j=1}^{n}\left[\left(y_t^{(i)}\right)^2 + \left(y_t^{(j)}\right)^2 - 2\left(\frac{y_t^{(i)} + y_t^{(j)}}{2}\right)^2\right] 1_{\{X_t=(i,j)\}} \\ &= \|Y_t\|_2^2 - \frac{1}{2}\sum_{i,j=1}^{n}\left(y_t^{(i)} - y_t^{(j)}\right)^2 1_{\{X_t=(i,j)\}}.\end{aligned}$$

En prenant les espérances conditionnelles par rapport à X_t et en utilisant le fait que X_t et Y_t sont indépendants, on obtient

$$\begin{aligned}\mathbb{E}\left(\|Y_{t+1}\|_2^4 \mid X_t = (i,j)\right) &= \mathbb{E}\left(\left(\|Y_{t+1}\|_2^2\right)^2 \mid X_t = (i,j)\right) \\ &= \mathbb{E}\left(\left[\|Y_t\|_2^2 - \frac{1}{2}\left(y_t^{(i)} - y_t^{(j)}\right)^2\right]^2\right) \\ &= \mathbb{E}\left(\|Y_t\|_2^4\right) - \mathbb{E}\left(\left(y_t^{(i)} - y_t^{(j)}\right)^2\|Y_t\|_2^2 - \frac{1}{4}\left(y_t^{(i)} - y_t^{(j)}\right)^4\right).\end{aligned}$$

En déconditionnant, on obtient

$$\begin{aligned}\mathbb{E}\left(\|Y_{t+1}\|_2^4\right) =& \mathbb{E}\left(\|Y_t\|_2^4\right) \\ &- \frac{1}{n(n-1)}\mathbb{E}\left[\|Y_t\|_2^2\sum_{i,j=1}^{n}\left(y_t^{(i)} - y_t^{(j)}\right)^2 - \frac{1}{4}\sum_{i,j=1}^{n}\left(y_t^{(i)} - y_t^{(j)}\right)^4\right].\end{aligned}$$

La première double somme peut s'écrire, en utilisant (5.8), comme

$$\sum_{i,j=1}^{n}\left(y_t^{(i)} - y_t^{(j)}\right)^2 = 2n\|Y_t\|_2^2.$$

De la même manière, en utilisant (5.8), la seconde double somme s'écrit

$$\begin{aligned}\sum_{i,j=1}^{n}\left(y_t^{(i)} - y_t^{(j)}\right)^4 &= \sum_{i,j=1}^{n}\left[\left(y_t^{(i)}\right)^4 + \left(y_t^{(j)}\right)^4 - 4\left(y_t^{(i)}\right)^3\left(y_t^{(j)}\right)\right. \\ &\qquad \left. - 4\left(y_t^{(i)}\right)\left(y_t^{(j)}\right)^3 + 6\left(y_t^{(i)}\right)^2\left(y_t^{(j)}\right)^2\right] \\ &= 2n\|Y_t\|_4^4 + 6\|Y_t\|_2^4.\end{aligned}$$

En mettant ensemble ces deux résultats, cela donne

$$\mathbb{E}\left(\|Y_{t+1}\|_2^4\right) = \mathbb{E}\left(\|Y_t\|_2^4\right) - \frac{1}{n(n-1)}\mathbb{E}\left[2n\|Y_t\|_2^4 - \frac{n}{2}\|Y_t\|_4^4 - \frac{3}{2}\|Y_t\|_2^4\right],$$

c'est-à-dire

$$\mathbb{E}\left(\|Y_{t+1}\|_2^4\right) = \left(1 - \frac{4n-3}{2n(n-1)}\right)\mathbb{E}\left(\|Y_t\|_2^4\right) + \frac{1}{2(n-1)}\mathbb{E}\left(\|Y_t\|_4^4\right),$$

ce qu'il fallait démontrer. ■

Nous sommes maintenant capables, en utilisant les théorèmes 5.1.4 et 5.1.5, d'obtenir des expressions explicites des moments d'ordre 4 de $\|Y_t\|_2$ et $\|Y_t\|_4$ en fonction de ceux de $\|Y_0\|_2$ et $\|Y_0\|_4$.

Corollaire 5.1.6 *Pour tout $n \geq 3$ et $t \geq 0$ nous avons,*

$$\mathbb{E}\left(\|Y_t\|_4^4\right) = \frac{6\alpha^t + n\beta^t}{n+6}\mathbb{E}\left(\|Y_0\|_4^4\right) + \frac{3(\beta^t - \alpha^t)}{n+6}\mathbb{E}\left(\|Y_0\|_2^4\right),$$
$$\mathbb{E}\left(\|Y_t\|_2^4\right) = \frac{2n(\beta^t - \alpha^t)}{n+6}\mathbb{E}\left(\|Y_0\|_4^4\right) + \frac{n\alpha^t + 6\beta^t}{n+6}\mathbb{E}\left(\|Y_0\|_2^4\right),$$

où

$$\alpha = 1 - \frac{2}{n-1} \text{ et } \beta = 1 - \frac{7n-6}{4n(n-1)}.$$

Preuve. Nous introduisons le vecteur colonne $U(t)$ défini par

$$U(t) = \begin{pmatrix} \mathbb{E}\left(\|Y_t\|_4^4\right) \\ \\ \mathbb{E}\left(\|Y_t\|_2^4\right) \end{pmatrix}.$$

Les relations (5.6) et (5.9) peuvent s'écrire, pour $t \geq 1$, $U(t) = AU(t-1)$, où A est la matrice 2×2 donnée par

$$A = \begin{pmatrix} 1 - \dfrac{7}{4(n-1)} & \dfrac{3}{4n(n-1)} \\ \\ \dfrac{1}{2(n-1)} & 1 - \dfrac{4n-3}{2n(n-1)} \end{pmatrix}.$$

Nous obtenons facilement, pour tout $t \geq 0$, $U(t) = A^t U(0)$. Les valeurs propres de A sont

$$\alpha = 1 - \frac{2}{n-1} \text{ et } \beta = 1 - \frac{7n-6}{4n(n-1)}.$$

Notons que comme $n \geq 3$, nous avons $0 \leq \alpha < \beta$. Les vecteurs propres V_α et V_β sont

$$V_\alpha = \begin{pmatrix} 1 \\ -n/3 \end{pmatrix} \text{ et } V_\beta = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Nous avons donc

$$A^t = \frac{1}{n+6}\begin{pmatrix} 6\alpha^t + n\beta^t & 3(\beta^t - \alpha^t) \\ 2n(\beta^t - \alpha^t) & n\alpha^t + 6\beta^t \end{pmatrix}.$$

Ce qui mène à

$$\mathbb{E}\left(\|Y_t\|_4^4\right) = \frac{6\alpha^t + n\beta^t}{n+6}\mathbb{E}\left(\|Y_0\|_4^4\right) + \frac{3(\beta^t - \alpha^t)}{n+6}\mathbb{E}\left(\|Y_0\|_2^4\right)$$

$$\mathbb{E}\left(\|Y_t\|_2^4\right) = \frac{2n(\beta^t - \alpha^t)}{n+6}\mathbb{E}\left(\|Y_0\|_4^4\right) + \frac{n\alpha^t + 6\beta^t}{n+6}\mathbb{E}\left(\|Y_0\|_2^4\right),$$

ce qu'il fallait démontrer. ■

Le théorème 5.1.3 fournit une expression exacte de $\mathbb{E}\left(\|Y_t\|_2^2\right)$. En utilisant ce résultat et le corollaire précédent, nous obtenons une expression de la variance de $\|Y_t\|_2^2$ que nous notons $\mathbb{V}\left(\|Y_t\|_2^2\right)$.

$$\begin{aligned}\mathbb{V}\left(\|Y_t\|_2^2\right) &= \mathbb{E}\left(\|Y_t\|_2^4\right) - \mathbb{E}\left(\|Y_t\|_2^2\right)^2 \\ &= \frac{2n(\beta^t - \alpha^t)}{n+6}\mathbb{E}\left(\|Y_0\|_4^4\right) + \frac{n\alpha^t + 6\beta^t}{n+6}\mathbb{E}\left(\|Y_0\|_2^4\right) - \left(1 - \frac{1}{n-1}\right)^{2t}\mathbb{E}\left(\|Y_0\|_2^2\right)^2.\end{aligned}$$

Par définition de $C_0^{(i)}$, utilisant le fait que $n = n_A + n_B$ et en introduisant la notation $p_{n,A} = n_A/n$, nous avons

$$\ell = \frac{1}{n}\sum_{i=1}^{n} C_0^{(i)} = \frac{1}{n}\left(n_A - n_B\right)m = (2p_{n,A} - 1)m. \tag{5.10}$$

Cette expression mène à

$$\begin{aligned}\|Y_0\|_2^2 &= \sum_{i=1}^{n}\left(y_0^{(i)}\right)^2 = \sum_{i=1}^{n}\left(C_0^{(i)}\right)^2 - n\ell^2 = nm^2 - n\ell^2 \\ &= 4nm^2 p_{n,A}(1 - p_{n,A}).\end{aligned} \tag{5.11}$$

Donc nous avons

$$\|Y_0\|_2^4 = 16n^2m^4p_{n,A}^2(1 - p_{n,A})^2. \tag{5.12}$$

En ce qui concerne la norme 4, nous avons

$$\begin{aligned}\|Y_0\|_4^4 &= \sum_{i=1}^{n}\left(y_0^{(i)}\right)^4 \\ &= \sum_{i=1}^{n}\left[\left(C_0^{(i)}\right)^4 - 4\left(C_0^{(i)}\right)^4\ell + 6\left(C_0^{(i)}\right)^2\ell^2 - 4C_0^{(i)}\ell^3 + \ell^4\right] \\ &= nm^4 - 4n\ell^2m^2 + 6n\ell^2m^2 - 4n\ell^4 + n\ell^4 \\ &= n(m - \ell)(m + \ell)(m^2 + 3\ell^2).\end{aligned}$$

En utilisant la relation (5.10), nous obtenons

$$\|Y_0\|_4^4 = 16nm^4 p_{n,A}(1-p_{n,A})(3p_{n,A}^2 - 3p_{n,A} + 1). \tag{5.13}$$

Notons que le maximum de la fonction $x(1-x)$ est atteint en $x = 1/2$ et sa valeur est égale à $1/4$. Le maximum de la fonction $x(1-x)(3x^2-3x+1)$ est atteint au point $x = (3 \pm \sqrt{3})/6$ et sa valeur est égale à $1/12$. Par conséquent

$$\|Y_0\|_2^2 \leq nm^2, \quad \|Y_0\|_2^4 \leq n^2m^4 \text{ et } \|Y_0\|_4^4 \leq 4nm^4/3. \tag{5.14}$$

Nous avons l'égalité quand $p_{n,A} = 1/2$ pour $\|Y_0\|_2^2$ et $\|Y_0\|_2^4$, et aussi quand $p_{n,A} = (3 \pm \sqrt{3})/6$ pour $\|Y_0\|_4^4$.

Théorème 5.1.7 *Pour tout $n \geq 3$, $\varepsilon > 0$, $\delta \in]0;1[$ et $t \geq \tau$, nous avons*

$$\mathbb{P}\{\|Y_t\|_\infty < \varepsilon\} \geq 1 - \delta,$$

où

$$\tau = \frac{4n(n-1)}{7n-6} \ln\left(\frac{13m^4 n}{3\varepsilon^4 \delta}\right).$$

Preuve. En utilisant les inégalités (5.14) et le corollaire 5.1.6, nous obtenons

$$\begin{aligned}
\mathbb{E}\left(\|Y_t\|_4^4\right) &\leq \frac{4nm^4(6\alpha^t + n\beta^t)}{3(n+6)} + \frac{3n^2m^4(\beta^t - \alpha^t)}{n+6} \\
&= \frac{1}{n+6}\left[\frac{13m^4n^2\beta^t}{3} - (3n-8)m^4 n\alpha^t\right] \\
&\leq \frac{13m^4n^2\beta^t}{3(n+6)} \leq \frac{13m^4 n\beta^t}{3}.
\end{aligned}$$

A partir du théorème 5.1.2 et l'inégalité de Markov, nous avons pour tout $\varepsilon > 0$ et $t \geq \tau$

$$\begin{aligned}
\mathbb{P}\{\|Y_t\|_\infty \geq \varepsilon\} = \mathbb{P}\{\|Y_t\|_\infty^4 \geq \varepsilon^4\} &\leq \mathbb{P}\{\|Y_t\|_4^4 \geq \varepsilon^4\} \\
&\leq \mathbb{P}\{\|Y_\tau\|_4^4 \geq \varepsilon^4\} \leq \frac{\mathbb{E}\left(\|Y_\tau\|_4^4\right)}{\varepsilon^4} \leq \frac{13m^4 n\beta^\tau}{3\varepsilon^4}.
\end{aligned}$$

Pour tout $x \in [0,1)$, nous avons $\ln(1-x) \leq -x$ ce qui est équivalent à $(1-x)^\tau \leq e^{-\tau x}$. Par définition de τ, cela mène à

$$\beta^\tau = \left(1 - \frac{7n-6}{4n(n-1)}\right)^\tau \leq e^{-\tau(7n-6)/(4n(n-1))} = \frac{3\varepsilon^4\delta}{13m^4 n}.$$

Nous obtenons donc pour $t \geq \tau$,

$$\mathbb{P}\{\|Y_t\|_\infty \geq \varepsilon\} \leq \delta, \text{ c'est-à-dire } \mathbb{P}\{\|Y_t\|_\infty < \varepsilon\} \geq 1 - \delta,$$

ce qu'il fallait démontrer. ∎

Le corollaire suivant montre qu'après τ interactions, chaque nœud est capable de fournir la proportion $\gamma_A = n_A/n$ de nœuds qui avaient initialement le symbole A, avec une probabilité d'au moins $1-\delta$, pour tout $\delta \in]0;1[$.

Corollaire 5.1.8 *Pour tout $n \geq 3$, $\varepsilon > 0$, $\delta \in]0;1[$ et $t \geq \tau$, nous avons*

$$\mathbb{P}\left\{\left|\frac{C_t^{(i)}+m}{2m} - \gamma_A\right| < \varepsilon, \text{ pour tout } i \in [\![1,n]\!]\right\} \geq 1-\delta,$$

où

$$\tau = \frac{4n(n-1)}{7n-6}\left(\ln(n) - \ln(48/13) - 4\ln(\varepsilon) - \ln(\delta)\right).$$

Preuve. En remplaçant ε par $2m\varepsilon$ dans le théorème 5.1.7, nous obtenons

$$\tau = \frac{4n(n-1)}{7n-6}\left(\ln(n) - \ln(48/13) - 4\ln(\varepsilon) - \ln(\delta)\right)$$

et par conséquent pour tout $t \geq \tau$, $\varepsilon > 0$ et $\delta \in]0;1[$, nous avons $\mathbb{P}\left\{\|Y_t\|_\infty \geq 2m\varepsilon\right\} \leq \delta$. Sachant que $Y_t = C_t - L$ et que $\ell = (2\gamma_A - 1)m$, nous obtenons

$$\begin{aligned}&\mathbb{P}\left\{\|C_t - L\|_\infty < 2m\varepsilon\right\} \geq 1-\delta\\ \iff &\mathbb{P}\left\{\left|\frac{C_t^{(i)}+m}{2m} - \gamma_A\right| < \varepsilon, \forall i \in [\![1,n]\!]\right\} \geq 1-\delta,\end{aligned}$$

ce qu'il fallait démontrer. ■

Le résultat suivant montre que, après τ interactions, chaque nœud est capable de fournir la valeur de n_A avec une probabilité supérieure à $1-\delta$, pour tout $\delta \in]0;1[$.

Corollaire 5.1.9 *Pour tout $n \geq 3$, $\delta \in]0;1[$ et $t \geq \tau$, nous avons*

$$\mathbb{P}\left\{\left\lfloor\frac{n\left(C_t^{(i)}+m\right)}{2m} + \frac{1}{2}\right\rfloor = n_A, \forall i \in [\![1,n]\!]\right\} \geq 1-\delta,$$

où

$$\tau = \frac{4n(n-1)}{7n-6}\left[5\ln(n) + \ln(104/3) - \ln(\delta)\right].$$

Preuve. En prenant $\varepsilon = m/n$ dans le théorème 5.1.7, nous obtenons

$$\tau = \frac{4n(n-1)}{7n-6}\left(5\ln(n) + \ln(104/3) - \ln(\delta)\right)$$

et par conséquent pour tout $t \geq \tau$ et $\delta \in]0,1[$, nous avons $\mathbb{P}\left\{\|Y_t\|_\infty \geq m/n\right\} \leq \delta$.

Puisque $Y_t = C_t - L$ et $\ell = (2n_A - n)m/n$, nous obtenons

$$\begin{aligned}
&\mathbb{P}\left\{\|C_t - L\|_\infty < m/n\right\} \geq 1-\delta \\
&\iff \mathbb{P}\left\{\left|C_t^{(i)} - \frac{(2n_A-n)m}{n}\right| < \frac{m}{n},\ \forall i \in [\![1,n]\!]\right\} \geq 1-\delta \\
&\iff \mathbb{P}\left\{\left|\frac{n\left(C_t^{(i)}+m\right)}{2m} - n_A\right| < \frac{1}{2},\ \forall i \in [\![1,n]\!]\right\} \geq 1-\delta \\
&\iff \mathbb{P}\left\{n_A < \frac{n\left(C_t^{(i)}+m\right)}{2m} + \frac{1}{2} < n_A + 1,\ \forall i \in [\![1,n]\!]\right\} \geq 1-\delta.
\end{aligned}$$

Cette dernière inégalité implique que

$$\mathbb{P}\left\{\left\lfloor\frac{n\left(C_t^{(i)}+m\right)}{2m} + \frac{1}{2}\right\rfloor = n_A, \forall i \in [\![1,n]\!]\right\} \geq 1-\delta,$$

ce qu'il fallait démontrer. ■

Le résultat suivant montre que, dans la mesure où $n_A \neq n_B$, après τ interactions, chaque nœud est capable de fournir l'indicatrice de $n_A > n_B$ avec une probabilité supérieure à $1-\delta$, pour tout $\delta \in]0,1[$, ce qui résout le problème de la majorité, tel que défini en section 3.4.

Corollaire 5.1.10 *Pour tout $n \geq 3$, $\delta \in]0;1[$ et $t \geq \tau$, en supposant que $n_A \neq n_B$, nous avons*

$$\mathbb{P}\left\{1_{\left\{C_t^{(i)}>0\right\}} = 1_{\{n_A>n_B\}}, \forall i \in [\![1,n]\!]\right\} \geq 1-\delta,$$

avec

$$\tau = \frac{4n(n-1)}{7n-6}\left[5\ln(n) + \ln(104/3) - \ln(\delta)\right].$$

Preuve. En prenant $\varepsilon = m/n$ dans le théorème 5.1.7, nous obtenons

$$\tau = \frac{4n(n-1)}{7n-6}\left(5\ln(n) + \ln(104/3) - \ln(\delta)\right)$$

et par conséquent, pour tout $t \geq \tau$ et $\delta \in]0,1[$, nous avons

$$\mathbb{P}\left\{\|C_t - L\|_\infty < m/n\right\} \geq 1-\delta.$$

D'autre part, $\ell = (n_A - n_B)m/n$, et puisque $n_A \neq n_B$, $|\ell| \geq m/n$, alors

$$\mathbb{P}\left\{1_{\left\{C_t^{(i)}>0\right\}} = 1_{\{\ell>0\}}, \forall i \in [\![1,n]\!]\right\} \geq 1-\delta,$$

puisque $1_{\{n_A>n_B\}} = 1_{\{\ell>0\}}$, nous obtenons

$$\mathbb{P}\left\{1_{\left\{C_t^{(i)}>0\right\}} = 1_{\{n_A>n_B\}}, \forall i \in [\![1,n]\!]\right\} \geq 1-\delta,$$

ce qu'il fallait démontrer. ■

5.1.4 Évaluation expérimentale

Cette section montre à quel point nos bornes théoriques sont proches des résultats de simulation. Nous comparons le résultat du corollaire 5.1.9 avec les résultats de simulations et nous comparons également ces bornes à celles obtenues par [65] ainsi que par nous-même [54]. Dans ces publications l'analyse est basée sur la norme 2. Les simulations concernent deux cas, traités séparément, pour le premier $n_A = n/2$ et pour le second $n_A = 3n/4$. Le problème du comptage est équivalent au problème de proportion avec $\varepsilon = 1/(2n)$. Par exemple, le problème du comptage avec $n = 1000$ (voir figure 5.2) est semblable au problème de proportion avec $\varepsilon = 0.0005$, l'avantage de ce problème est qu'il peut être comparé aux résultats obtenus dans [54].

Une simulation se déroule en plusieurs étapes. En premier lieu, n_A nœuds sont initialisés à m et n_B nœuds sont initialisés à $-m$ (sans perte de généralité, nous prenons $m = 1$). Ensuite, à chaque étape de la simulation, deux nœuds sont choisis au hasard pour interagir et mettre à jour leur état. La simulation s'arrête lorsque tous les nœuds sont capables de connaître n_A.

Les simulations ont été effectuées avec deux paramétrages différents : dans le premier, $n_A = n/2$ et dans le second, $n_A = 3n/4$.

Pour chaque paramétrage, nous avons exécuté N simulations indépendantes et avons mémorisé et ordonné le nombre d'interactions effectuées pour arriver à la convergence $\tau_1 \leq \tau_2 \leq \ldots \leq \tau_N$. Le temps de convergence est alors $\tau_{\lceil N(1-\delta)\rceil}$, avec $\delta \in]0, 1[$.

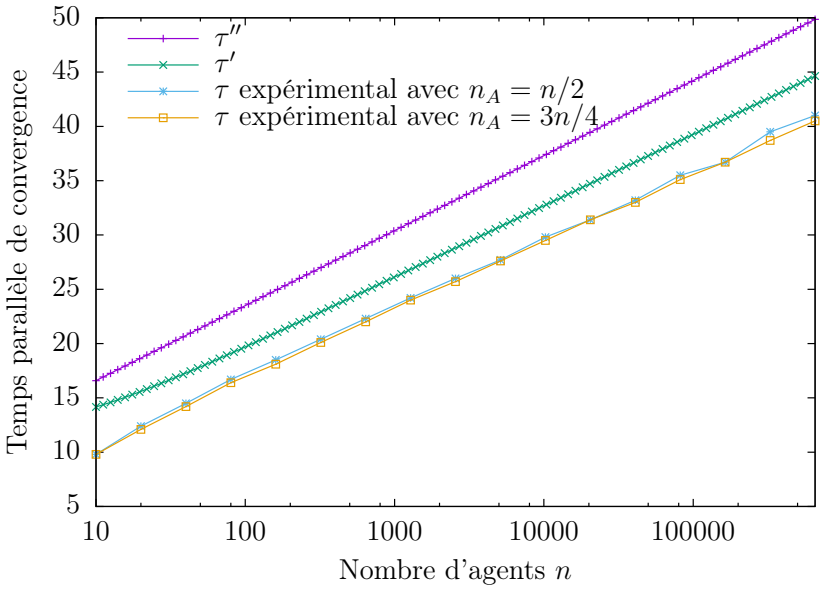


FIGURE 5.1 : Temps parallèle de convergence en fonction de n quand $\delta = 10^{-3}$ et $N = 10^5$ pour $n < 20000$, $N = 10^4$ pour $n > 20000$.

Les figures 5.1 et 5.2 donnent le temps parallèle de convergence (rappelons que le temps parallèle de convergence est le temps de convergence divisé par n) $\tau_{\lceil N(1-\delta)\rceil}/n$ pour chacun des deux scénarios, la borne τ' obtenue à partir de la relation (5.1.9), et la borne τ'' obtenue à partir du théorème 4 de [54], c'est-à-dire

$\tau'' = 4\ln 2 + 3\ln n - \ln \delta$. La figure 5.1 montre clairement que la borne τ' donnée par le corollaire 5.1.9 est très proche du temps parallèle obtenu par les simulations, et améliore considérablement les résultats précédents, c'est-à-dire τ''. La figure 5.2 confirme la finesse de notre analyse théorique.

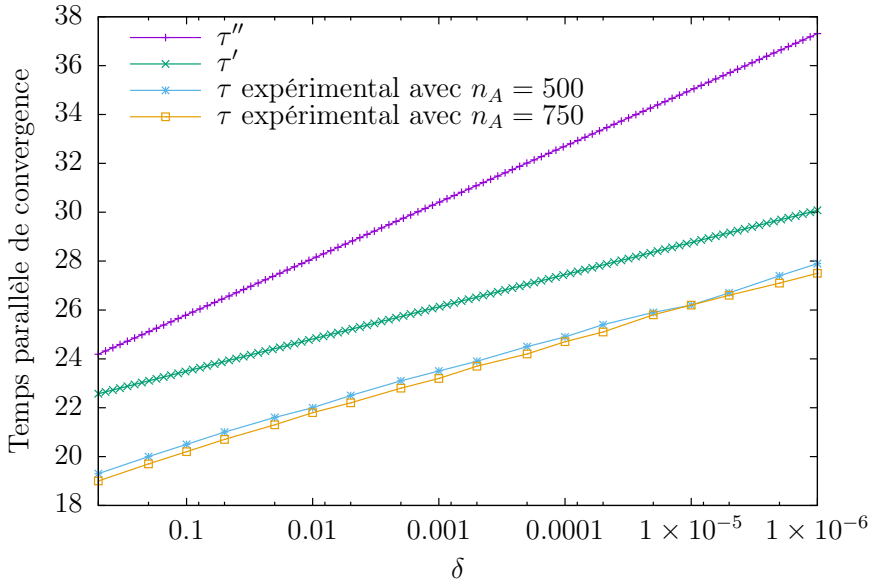


FIGURE 5.2 : Temps parallèle de convergence en fonction de δ quand $n = 10^3$ et $N = 10^6$.

5.1.5 Conclusion

Dans cette section, nous avons présenté une analyse très précise du temps nécessaire à chaque nœud pour résoudre les problèmes de comptage et de proportion. Notre travail repose sur l'utilisation de la norme 4. Une comparaison des limites obtenues à la fois avec la norme 4 et la norme 2, montre à quel point il est avantageux d'utiliser la norme 4 plutôt que la norme 2. On pourrait penser que l'utilisation d'une norme d, pour $d > 4$, donnerait des résultats plus affinés mais cela conduirait à une analyse beaucoup plus complexe.

5.2 Protocoles de moyenne avec des entiers

Cette section aborde le protocole de moyenne sous l'angle du problème de la proportion avec des entiers. Le point remarquable de cette section est la démonstration qu'un protocole basé sur la moyenne avec des entiers converge en $O(\log n)$ (temps parallèle) vers un état où la différence entre deux valeurs du système est toujours inférieure où égale à 2. Cela nous permet de prouver, en section 5.2.4, le caractère optimal de notre protocole. Nous abordons aussi le problème de l'influence de la partie fractionnaire de la moyenne générale (qui reste invariante) sur le temps de convergence, autant du point de vue de la démonstration (théorème 5.2.10) que des expérimentations (section 5.2.6), ce dernier point était jusque-là inédit.

Cette section reprend des résultats de trois de nos publications [54, 55, 59].

5.2.1 Introduction

Comme cela a été montré en section 3.4.2, le problème de la majorité a été très étudié. Dans ce chapitre, nous nous concentrons sur une question connexe à celui-ci et plus générale. Ainsi, au lieu de demander à chaque agent de répondre "oui" si une majorité d'agents commence son exécution dans l'état A, on peut se demander s'il est possible, pour chaque agent, de calculer *rapidement* et avec une précision élevée, la proportion d'agents qui a commencé dans l'état A. Répondre à une telle question est très important. Par exemple dans le contexte de la surveillance des maladies infectieuses animales, différents types d'alertes sont déclenchés selon la proportion de la population infectée (c'est-à-dire, Alerte 1 est déclenchée si moins de 0.05% de la population est infectée, Alerte 2 si cette proportion se trouve dans $]0.05\%, 3.0\%]$, Alerte 3 si elle se trouve dans $]3.0\%, 10\%]$, et ainsi de suite).

Nous répondons par l'affirmative à cette question, et nous proposons un protocole de population qui permet à chaque agent de converger vers un état à partir duquel, lorsqu'il est interrogé, il fournit la proportion d'agents qui se trouvent dans un état donné.

Plus précisément, chaque agent est une machine à $(2m + 1)$ états, $m \geq 1$, où m est la valeur initiale associée à l'état A et $-m$ est associée à l'état B. Chaque agent commence son exécution avec m ou $-m$, et chaque paire d'agents qui se rencontrent adopte la moyenne de leurs états (ou aussi proche qu'elle peut l'être du fait que les états sont restreints à des entiers, comme cela sera précisé dans la section 5.2.2). Cette méthode préserve la somme des états initiaux, et après $O(\log n)$ interactions, chaque agent converge avec une probabilité élevée vers un état à partir duquel il peut calculer la proportion d'agents qui ont démarré dans un état donné, et la précision ε du résultat est en $O(1/m)$.

Plus précisément, notre protocole garantit que chaque agent est capable de calculer la proportion d'agents qui ont démarré dans un état spécifique avec une précision inférieure à ε en utilisant $(3/(2\varepsilon)) + 1$ états, avec bien entendu $\varepsilon \in]0; 1[$.

Ceci est réalisé en moins de $n\,(3.12 \ln n - 2 \ln \varepsilon - 6.59 \ln \delta + 1.88)$ interactions avec une probabilité supérieure à $1 - \delta$, pour tout $\delta \in]0; 1[$.

Le problème du comptage, quant à lui, généralise le problème de la proportion. Il s'agit pour chaque agent, de converger vers un état où chaque agent est capable de calculer n_A et d'en déduire n_B, où n_A et n_B représentent respectivement le nombre d'agents qui ont commencé dans l'état A et B.

En s'appuyant sur les preuves présentées pour la proportion, et en modifiant légèrement la fonction de sortie, en supposant que n est connu, nous montrons dans la section 5.2.4 que nous résolvons le problème de comptage avec $O(n \log n)$ interactions et avec seulement $O(n)$ états par agent. Nous montrons également que tout protocole résolvant le problème de comptage nécessite $\Omega(\log n)$ en temps parallèle pour converger et $\Omega(n)$ états. De ce fait, nous prouvons que notre algorithme est une solution optimale en espace et en temps pour résoudre à la fois le problème du comptage et celui de la proportion.

Le reste de cette section est organisé comme suit. Le protocole pour calculer la

proportion est présenté dans la section 5.2.2. Une analyse du protocole est proposée dans la section 5.2.3. Nous montrons dans la section 5.2.4, que notre protocole est optimal à la fois en espace et en temps. Nous avons simulé notre protocole pour illustrer notre analyse théorique en section 5.2.6. Finalement, la section 5.2.7 conclut.

5.2.2 Calculer la proportion

Notre protocole utilise la technique de la moyenne pour calculer la proportion d'agents qui ont commencé leur exécution dans l'état A. Nous notons $C_t^{(i)}$ l'état de l'agent i à l'instant t. Nous nommons le processus stochastique $C = \{C_t,\ t \geq 0\}$ où $C_t = \left(C_t^{(1)}, \ldots, C_t^{(n)}\right)$ représente la configuration du système à l'instant t. L'ensemble d'entrée du protocole est $\Sigma = \{A, B\}$, et la fonction d'entrée ι est définie par $\iota(A) = m$ et $\iota(B) = -m$, m étant un entier strictement positif que l'on déterminera. Cela signifie que, pour tout $i = [\![1, n]\!]$, nous avons $C_0^{(i)} \in \{-m, m\}$. A chaque instant discret t, deux indices distincts i et j sont choisis parmi $1, \ldots, n$ avec probabilité $p_{i,j}(t)$. Une fois choisi, le couple (i, j) interagit, et les deux agents mettent à jour leur état local respectif $C_t^{(i)}$ et $C_t^{(j)}$ en appliquant la fonction de transition f, de la façon suivante

$$\left(C_{t+1}^{(i)}, C_{t+1}^{(j)}\right) = f(C_t^{(i)}, C_t^{(j)}) = \left(\left\lfloor \frac{C_t^{(i)} + C_t^{(j)}}{2} \right\rfloor, \left\lceil \frac{C_t^{(i)} + C_t^{(j)}}{2} \right\rceil\right) \quad (5.15)$$
$$\text{et} \quad C_{t+1}^{(r)} = C_t^{(r)} \text{ pour } r \neq i, j.$$

L'ensemble Q des états est $[\![-m, m]\!]$. La fonction de sortie est définie, pour tout $x \in Q$, par

$$\omega_A(x) = (m + x)/2m.$$

Par conséquent, l'ensemble de sortie Y est l'ensemble de toutes les valeurs possibles de ω_A, c'est-à-dire

$$Y = \left\{0, \frac{1}{2m}, \frac{2}{2m}, \ldots, \frac{2m-2}{2m}, \frac{2m-1}{2m}, 1\right\}.$$

5.2.3 Analyse du protocole de proportion

Nous notons X_t la variable aléatoire représentant le choix à l'instant t de deux indices distincts i et j parmi $1, \ldots, n$ avec probabilité $p_{i,j}(t)$, c'est-à-dire $\mathbb{P}\{X_t = (i, j)\} = p_{i,j}(t)$. Nous supposons que la suite $\{X_t,\ t \geq 0\}$ est une suite de variables aléatoires indépendantes et identiquement distribuées. Puisque C_t est entièrement déterminée par les valeurs de $C_0, X_0, X_1, \ldots, X_{t-1}$, cela signifie en particulier que les variables aléatoires X_t et C_t sont indépendantes et que le processus stochastique C, représentant la suite des configurations, est une chaîne de Markov homogène en temps discret. Nous supposons que X_t est uniformément distribuée, c'est-à-dire

$$\mathbb{P}\{X_t = (i, j)\} = p_{i,j}(t) = \frac{1_{\{i \neq j\}}}{n(n-1)}.$$

Nous utliserons par la suite la norme euclidienne notée $\|.\|$ et la norme infini notée $\|.\|_\infty$ définie pour tout $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ par

$$\|x\| = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2} \text{ et } \|x\|_\infty = \max_{i=1,\ldots,n} |x_i|.$$

Il est bien connu que ces normes satisfont $\|x\|_\infty \leq \|x\| \leq \sqrt{n}\|x\|_\infty$.

Lemme 5.2.1 *Pour tout $t \geq 0$, nous avons*

$$\sum_{i=1}^{n} C_t^{(i)} = \sum_{i=1}^{n} C_0^{(i)}.$$

Preuve. La preuve est immédiate puisque la transformation de C_t à C_{t+1} décrite dans la relation (5.15) ne change pas la somme des composantes de C_{t+1}. En effet, à partir de la relation (5.15), nous avons $C_{t+1}^{(i)} + C_{t+1}^{(j)} = C_t^{(i)} + C_t^{(j)}$ et les autres entrées ne changent pas leurs valeurs. ■

Nous notons ℓ la moyenne de la somme des composantes de C_t et L le vecteur ligne de $\mathbb{R}^n$ avec toutes ses composantes égales à ℓ, c'est-à-dire

$$\ell = \frac{1}{n} \sum_{i=1}^{n} C_t^{(i)} \text{ et } L = (\ell, \ldots, \ell).$$

Notre analyse est organisée comme suit. Par un calcul sur la différence $\|C_{t+1} - L\|^2 - \|C_t - L\|^2$, le corollaire 5.2.3 établit une récurrence qui montre que $\mathbb{E}(\|C_t - L\|)$ est bornée par une suite arithmético-géométrique (c'est à dire dont la formule de récurrence a la forme $u_{n+1} = au_n + b$), le théorème 5.2.4 nous fournit une expression générale de la borne. Cette formulation permet de montrer facilement par une application directe de l'inégalité de Markov que dans le cas où la moyenne des entrées de C_0 a une partie fractionnaire égale à 0.5 ($\ell - \lfloor \ell \rfloor = 0.5$), le protocole converge en $O(\log n)$, avec une probabilité élevée, vers un état où l'écart maximal entre deux composantes du vecteur C_t est égal à 2. Pour les autres cas nous ne pouvons pas appliquer l'inégalité de Markov, nous montrerons cependant la même chose mais par un cheminement un peu plus complexe. En nous appuyant sur le théorème 5.2.4, nous montrons dans le théorème 5.2.9 que le processus stochastique C_t appartient à la boule de rayon $\sqrt{\rho n}$ et de centre L, avec une probabilité élevée et un temps parallèle de $O(\log n)$, pour $\rho \in]1/4; 5/4[$.

Ensuite, en supposant que C_t est dans la boule de rayon $\sqrt{\rho n}$ et que $\ell - \lfloor \ell \rfloor \neq 1/2$, nous démontrons que C_t converge vers un état où la distance maximale entre deux noeuds est égale à 2, avec une probabilité élevée et un temps parallèle de $O(\log n)$ (théorème 5.2.11).

Puis, si $\ell - \lfloor \ell \rfloor = 1/2$, nous appliquons le théorème 5.2.8, et si $\ell - \lfloor \ell \rfloor \neq 1/2$, nous appliquons successivement les théorèmes 5.2.9 et 5.2.11. Pour prouver notre théorème principal 5.2.12 qui montre que dans C_t converge vers une configuration où la différence maximale entre deux nœuds est égale 2, avec une probabilité élevée et un temps parallèle de $O(\log n)$.

Enfin, nous avons tous les outils nécessaires pour construire une fonction de sortie qui résout le problème de proportion avec une précision de ε, une probabilité de $1-\delta$, un temps parallèle de $O(\log n - \log\varepsilon - \log\delta)$, et avec $O(1/\varepsilon)$ états. Ceci est démontré dans le théorème 5.2.13.

Il faut noter que dans ce manuscrit, contrairement à ce qui a été fait dans nos publications [54, 55, 59], les théorèmes 5.2.9 et 5.2.11 sont paramétrés. Ceci nous permet, dans le théorème 5.2.12, d'optimiser les coefficients.

Lemme 5.2.2 *Soient $t, s \in \mathbb{N}$ avec $t \geq s$, $y \in \mathbb{R}$ et sous l'hypothèse que $\mathbb{P}\{Y_s \geq y\} \neq 0$, alors*

$$\mathbb{E}(Y_{t+1} \mid Y_s \geq y) \leq \left(1 - \frac{1}{n-1}\right)\mathbb{E}(Y_t \mid Y_s \geq y) + \frac{n}{4(n-1)}, \tag{5.16}$$

où Y_s est défini par $Y_s = \|C_s - L\|^2$.

Preuve. À partir de la relation (5.15), nous avons, pour tout $t \geq 0$,

$$Y_{t+1} = Y_t - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\left(C_t^{(i)} - C_t^{(j)}\right)^2 - 1_{\{C_t^{(i)}+C_t^{(j)} \text{ odd}\}}\right]1_{\{X_t=(i,j)\}}. \tag{5.17}$$

En multipliant les deux membres de l'égalité par $1_{\{Y_s\geq y\}}$, nous obtenons

$$\begin{aligned} Y_{t+1}1_{\{Y_s\geq y\}} &= Y_t 1_{\{Y_s\geq y\}} \\ &\quad - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\left(C_t^{(i)} - C_t^{(j)}\right)^2 - 1_{\{C_t^{(i)}+C_t^{(j)} \text{ odd}\}}\right]1_{\{Y_s\geq y\}}1_{\{X_t=(i,j)\}}. \end{aligned}$$

En prenant l'espérance, en utilisant le fait que X_t et C_t sont indépendants, nous avons

$$\begin{aligned} \mathbb{E}\left(Y_{t+1}1_{\{Y_s\geq y\}}\right) &= \mathbb{E}\left(Y_t 1_{\{Y_s\geq y\}}\right) \\ &\quad - \frac{1}{2}\mathbb{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\left(C_t^{(i)} - C_t^{(j)}\right)^2 - 1_{\{C_t^{(i)}+C_t^{(j)} \text{ odd}\}}\right]1_{\{Y_s\geq y\}}\right)p_{i,j}(t). \end{aligned}$$

Puisque

$$p_{i,j}(t) = \frac{1}{n(n-1)},$$

nous pouvons écrire

$$\begin{aligned} \mathbb{E}\left(Y_{t+1}1_{\{Y_s\geq y\}}\right) &= \mathbb{E}\left(Y_t 1_{\{Y_s\geq y\}}\right) \\ &\quad - \frac{1}{2n(n-1)}\mathbb{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\left(C_t^{(i)} - C_t^{(j)}\right)^2 - 1_{\{C_t^{(i)}+C_t^{(j)} \text{ odd}\}}\right]1_{\{Y_s\geq y\}}\right). \end{aligned}$$

De manière similaire à la démonstration du théorème 5.1.3, pour les réels, nous avons

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(C_t^{(i)}-C_t^{(j)}\right)^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\left(C_t^{(i)}-\ell\right)-\left(C_t^{(j)}-\ell\right)\right)^2\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\Big[\left(C_t^{(i)}-\ell\right)^2+\left(C_t^{(j)}-\ell\right)^2\\ &\qquad -2\left(C_t^{(i)}-\ell\right)\left(C_t^{(j)}-\ell\right)\Big]\\ &= 2n\|C_t-L\|^2-2\sum_{i=1}^{n}\left(C_t^{(i)}-\ell\right)\sum_{j=1}^{n}\left(C_t^{(j)}-\ell\right)\\ &= 2n\|C_t-L\|^2=2nY_t,\end{aligned}$$

q_t étant le nombre de composantes impaires de C_t, nous avons

$$\sum_{i=1}^{n}\sum_{j=1}^{n}1_{\{C_t^{(i)}+C_t^{(j)}\text{ odd}\}}=2q_t(n-q_t).$$

La fonction g, définie pour $x\in[0,n]$ par $g(x)=x(n-x)$, possède un maximum au point $x=n/2$, par conséquent $0\le g(x)\le n^2/4$. Cela donne

$$\sum_{i=1}^{n}\sum_{j=1}^{n}1_{\{C_t^{(i)}+C_t^{(j)}\text{ odd}\}}\le\frac{n^2}{2}.$$

Il s'ensuit que

$$\mathbb{E}\left(Y_{t+1}1_{\{Y_s\ge y\}}\right)\le\left(1-\frac{1}{n-1}\right)\mathbb{E}\left(Y_t1_{\{Y_s\ge y\}}\right)+\frac{n}{4(n-1)}\mathbb{P}\{Y_s\ge y\},$$

et puisque $\mathbb{P}\{Y_s\ge y\}\neq 0$, nous avons

$$\mathbb{E}\left(Y_t\mid Y_s\ge y\right)\le\left(1-\frac{1}{n-1}\right)\mathbb{E}\left(Y_s\mid Y_s\ge y\right)+\frac{n}{4(n-1)},$$

ce qu'il fallait démontrer. ■

Corollaire 5.2.3 *Soit $t\ge 0$, nous avons*

$$\mathbb{E}\left(Y_{t+1}\right)\le\left(1-\frac{1}{n-1}\right)\mathbb{E}\left(Y_t\right)+\frac{n}{4(n-1)}. \tag{5.18}$$

Preuve. La preuve est immédiate à partir du lemme 5.2.2 précédent, en prenant $s=0$ et $y=0$, dans ce cas $\mathbb{P}\{Y_s\ge y\}=1$. ■

Théorème 5.2.4 *Pour tout $t\ge 0$, nous avons*

$$\mathbb{E}\left(Y_t\right)\le\left(1-\frac{1}{n-1}\right)^t\mathbb{E}\left(Y_0\right)+\frac{n}{4}. \tag{5.19}$$

Preuve. D'après le corollaire 5.2.3, $\mathbb{E}(Y_t)$ est une suite définie par récurrence. Nous allons prouver l'exactitude de (5.19) par récurrence.

L'inégalité (5.19) est évidemment vraie pour $t = 0$.

Supposons qu'elle soit vraie pour t et appliquons le corollaire 5.2.3, cela donne

$$\begin{aligned}\mathbb{E}\left(Y_{t+1}\right) &\leq \left(1-\frac{1}{n-1}\right)\left[\left(1-\frac{1}{n-1}\right)^{t}\mathbb{E}\left(Y_0\right)+\frac{n}{4}\right]+\frac{n}{4(n-1)}\\ &= \left(1-\frac{1}{n-1}\right)^{t+1}\mathbb{E}(Y_0)+\frac{n}{4}.\end{aligned}$$

La récurrence est établie, ce qui termine la preuve. ■

Voici une version conditionnelle de ce théorème.

Théorème 5.2.5 *Soient $t, s \in \mathbb{N}$ avec $t \geq s \geq 0$, $y \in \mathbb{R}$ et sous l'hypothèse que $\mathbb{P}\{Y_s \geq y\} \neq 0$, alors*

$$\mathbb{E}\left(Y_t \mid Y_s \geq y\right) \leq \left(1-\frac{1}{n-1}\right)^{t-s}\mathbb{E}\left(Y_s \mid Y_s \geq y\right)+\frac{n}{4}. \tag{5.20}$$

Preuve. De manière analogue au théorème précédent, d'après le lemme 5.2.2, $\mathbb{E}(Y_t \mid Y_s \geq y)$ est une suite définie par récurrence.

L'inégalité (5.20) est évidemment vraie pour $t = s$.

Supposons l'inégalité (5.20) vraie pour $t \geq s$, et appliquons le lemme 5.2.2, cela donne

$$\begin{aligned}\mathbb{E}\left(Y_{t+1} \mid Y_s \geq y\right) &\leq \left(1-\frac{1}{n-1}\right)\left[\left(1-\frac{1}{n-1}\right)^{t-s}\mathbb{E}\left(Y_s \mid Y_s \geq y\right)+\frac{n}{4}\right]+\frac{n}{4(n-1)}\\ &= \left(1-\frac{1}{n-1}\right)^{t+1-s}\mathbb{E}\left(Y_s \mid Y_s \geq y\right)+\frac{n}{4}.\end{aligned}$$

La récurrence est établie, ce qui termine la preuve. ■

Voici un lemme qui prouve la décroissance de (Y_t). Ce lemme sera utilisé pour la démonstration du théorème 5.2.9.

Lemme 5.2.6 *La suite $Y_t = \|C_t - L\|^2$ est décroissante.*

Preuve. La preuve est immédiate à partir de l'égalité (5.17). Si $C_t^{(i)} + C_t^{(j)}$ est impair, alors $C_t^{(i)} \neq C_t^{(j)}$, donc, dans tous les cas, nous avons

$$\left(C_t^{(i)} - C_t^{(j)}\right)^2 - 1_{\{C_t^{(i)}+C_t^{(j)} \text{ impair}\}} \geq 0.$$

Nous avons donc bien $Y_{t+1} \leq Y_t$. ■

Le lemme suivant prouve la décroissance de $(\|C_t - L\|_\infty)_{t\in\mathbb{N}}$. Ce lemme est utilisé pour la démonstration du lemme de croissance de la convergence 5.2.14 qui est essentiel dans le chapitre 7 sur la détection de convergence.

Lemme 5.2.7 *La suite $(\|C_t - L\|_\infty)_{t\in\mathbb{N}}$ est décroissante.*

Preuve. À chaque instant $t \geq 0$, nous avons

$$\ell - \|C_t - L\|_\infty \leq C_t^{(i)} \leq \ell + \|C_t - L\|_\infty, \text{ pour tout } i \in [\![1, n]\!]. \tag{5.21}$$

Suite à une interaction à l'instant $t \geq 0$, à partir de (5.15), nous avons

$$\begin{cases} \min\{C_t^{(i)}, C_t^{(j)}\} \leq C_{t+1}^{(i)} \leq C_{t+1}^{(j)} \leq \max\{C_t^{(i)}, C_t^{(j)}\}, \text{ sachant que } X_t = (i,j) \\ C_{t+1}^{(r)} = C_t^{(r)}, \forall r \in [\![1, n]\!] \setminus \{i, j\}, \text{ sachant que } X_t = (i,j), \end{cases}$$

donc, pour tout $t \geq 0$,

$$\min_{1\leq r\leq n} C_t^{(r)} \leq C_{t+1}^{(i)} \leq \max_{1\leq r\leq n} C_t^{(r)}, \text{ pour tout } i \in [\![1, n]\!].$$

En combinant avec (5.21), à chaque instant $t \geq 0$ et pour tout $i \in [\![1, n]\!]$, nous obtenons

$$\ell - \|C_t - L\|_\infty \leq \min_{1\leq r\leq n} C_t^{(r)} \leq C_{t+1}^{(i)} \leq \max_{1\leq r\leq n} C_t^{(r)} \leq \ell + \|C_t - L\|_\infty$$
$$\ell - \|C_t - L\|_\infty \leq C_{t+1}^{(i)} \leq \ell + \|C_t - L\|_\infty,$$

ce qui peut s'écrire

$$\ell - C_{t+1}^{(i)} \leq \|C_t - L\|_\infty \text{ et } C_{t+1}^{(i)} - \ell \leq \|C_t - L\|_\infty, \text{ pour tout } i \in [\![1, n]\!],$$

par conséquent

$$\|C_{t+1} - L\|_\infty \leq \|C_t - L\|_\infty,$$

ce qu'il fallait démontrer. ■

Nous allons d'abord traiter un cas particulier où la moyenne de la somme a une partie fractionnaire égale à 0.5. Dans ce cas, et dans ce cas seulement, nous pouvons appliquer l'inégalité de Markov, et obtenir une formule avec des coefficients assez faibles.

Théorème 5.2.8 *Pour tout $\delta \in]0; 1[$, si $\ell - \lfloor \ell \rfloor = 0.5$ et s'il existe une constante K telle que $\|C_0 - L\|_\infty \leq K$, alors, pour tout $t \geq (n-1)\,(2\ln K + \ln n - \ln \delta - \ln 2)$, nous avons*

$$\mathbb{P}\{\|C_t - L\|_\infty \neq 1/2\} \leq \delta.$$

Preuve. Voir annexe section B.1. ■

Voici un théorème qui donne une formule explicite du temps nécessaire pour que le vecteur C_t se rapproche du vecteur L au point d'appartenir à la boule de rayon $\sqrt{\rho n}$ et de centre L, ceci avec une probabilité supérieure à $1 - \delta$, pour $\delta \in]0; 1[$ et $\rho > 1/4$. μ est un paramètre du théorème que nous fixerons ultérieurement afin d'optimiser les coefficients.

Théorème 5.2.9 *Soient* $\delta \in]0;1[$, $\rho > 1/4$ *et* $\mu > \max\{4/(4\rho-1), 4(e-1)\}$. *S'il existe* $K \geq \max\left\{\sqrt{\rho n}, \|C_0 - L\|\right\}$, *alors pour tout* $t \geq n\theta$, *nous avons*

$$\mathbb{P}\left\{\|C_t - L\|^2 < \rho n\right\} \geq 1 - \delta,$$

où

$$\theta = 2\ln K - \ln n + \ln \mu - (\ln(\rho\mu)/[\ln(4\rho\mu) - \ln(\mu+4)])\ln \delta.$$

Preuve. Voir annexe section B.1 ■

Le théorème suivant nous donne une formule explicite du temps pour que, en partant d'une configuration où le vecteur C_t est dans la boule de rayon $\sqrt{\rho n}$ et de centre L, nous parvenons à une configuration où l'écart maximal entre deux composantes de C_t soit 2, ceci avec une probabilité supérieure à $1-\delta$, sachant que $\delta \in]0;1[$ et $\rho \in]1/4;5/4[$. Ce théorème introduit la donnée λ dont la valeur absolue est la distance de la moyenne ℓ des composantes du vecteur C_t avec l'entier le plus proche, plus précisément $\lambda = \ell - \lceil \ell - 0.5 \rceil$.

Théorème 5.2.10 *Soit* $\delta \in]0;1[$ *et* $\rho \in]1/4, 5/4[$, *si* $\|C_0 - L\| \leq \sqrt{\rho n}$, $\ell - \lfloor \ell \rfloor \neq \frac{1}{2}$ *et* $\lambda = \ell - \lceil \ell - 0.5 \rceil$ *alors nous avons, pour tout* $t \geq \tau$,

$$\mathbb{P}\{\|C_t - L\|_\infty \geq 3/2\} \leq \delta,$$

où

$$\tau = \frac{25(n-1)}{12(2-\rho-\lambda^2-|\lambda|)}\left(\ln n - \ln \delta - 2\ln 2 + \ln\left(\rho + \lambda^2\right)\right).$$

Preuve. Voir annexe section B.1 ■

Ce théorème est une version plus générale du théorème précédent. Il permet d'avoir une majoration du temps de convergence avec probabilité élevée, sans avoir à connaître au préalable λ.

Théorème 5.2.11 *Soit* $\delta \in]0;1[$ *et* $\rho \in]1/4;5/4[$. *Si* $\|C_0 - L\| \leq \sqrt{\rho n}$ *et* $\ell - \lfloor \ell \rfloor \neq 1/2$ *alors nous avons pour tout* $t \geq \tau$,

$$\mathbb{P}\{\|C_t - L\|_\infty \geq 3/2\} \leq \delta,$$

où

$$\tau = \frac{25(n-1)}{3(5-4\rho)}\left(\ln n - \ln \delta - 4\ln 2 + \ln\left(4\rho + 1\right)\right).$$

Preuve. Dans le théorème précédent, si on considère τ comme une fonction de $|\lambda| \in [0;1/2]$, τ est une fonction croissante dont le maximum est atteint quand $|\lambda| = 1/2$. Ne connaissant pas λ, nous pouvons obtenir la borne supérieure de τ en prenant la borne supérieure de $|\lambda|$, c'est-à-dire $1/2$. Cela nous permet d'avoir un minorant de t. Nous avons donc

$$\begin{aligned}
\tau &= \frac{25(n-1)}{12(2-\rho-1/4-1/2)}\left(\ln n - \ln \delta - 2\ln 2 + \ln\left(\rho + 1/4\right)\right) \\
&= \frac{25(n-1)}{12(5/4-\rho)}\left(\ln n - \ln \delta - 4\ln 2 + \ln\left(4\rho + 1\right)\right) \\
&= \frac{25(n-1)}{3(5-4\rho)}\left(\ln n - \ln \delta - 4\ln 2 + \ln\left(4\rho + 1\right)\right),
\end{aligned}$$

ce qu'il fallait démontrer. ■

Nous présentons maintenant un théorème, synthèse des théorèmes 5.2.9 et 5.2.11. Ce théorème optimise les paramètres pour donner un résultat numérique du temps nécessaire pour que le protocole basé sur la moyenne avec des entiers arrive à un état où la différence maximum entre deux états soit deux, ceci avec une probabilité supérieure à $1-\delta$.

Théorème 5.2.12 *Pour tout $\delta \in]0;1[$, s'il existe une constante K telle que $\|C_0 - L\| \leq K$, alors, pour tout $t \geq n\,(2\ln K + 2.12\ln n - 6.59\ln\delta + 1.88)$, nous avons*

$$\mathbb{P}\left\{\|C_t - L\|_\infty \geq 3/2\right\} \leq \delta \Longleftrightarrow \mathbb{P}\{\|C_t - L\|_\infty < 3/2\} \geq 1-\delta.$$

Preuve. Voir annexe section B.1 ■

Nous appliquons ces résultats pour calculer la proportion γ_A d'agents dont l'entrée initiale est A, et où $\gamma_A = n_A/(n_A+n_B) = n_A/n$. Rappelons que la fonction ω_A est définie, pour tout $x \in Q$, par

$$\omega_A(x) = (m+x)/(2m).$$

Théorème 5.2.13 *Pour tout $\delta \in]0,1[$ et pour $\varepsilon \in]0,1[$, en prenant $m = \lceil 3/(4\varepsilon)\rceil$, nous avons, pour tout $t \geq \tau$,*

$$\mathbb{P}\left\{\left|\omega_A\left(C_t^{(i)}\right) - \gamma_A\right| < \varepsilon \text{ pour tout } i \in [\![1,n]\!]\right\} \geq 1-\delta,$$

où

$$\tau = n\,(3.12\ln n - 2\ln\varepsilon - 6.59\ln\delta + 1.88)\,.$$

Preuve. Nous avons $\|C_0 - L\| \leq m\sqrt{n}$. En appliquant le théorème 5.2.12, avec $K = \sqrt{n}/\varepsilon \geq \lceil 3/(4\varepsilon)\rceil\sqrt{n} = m\sqrt{n}$, nous obtenons, pour tout $\delta \in]0;1[$
et $t \geq n\,(3.12\ln n - 2\ln\varepsilon - 6.59\ln\delta + 1.88)$,

$$\mathbb{P}\{\|C_t - L\|_\infty \geq 3/2\} \leq \delta$$

ou de manière équivalente

$$\mathbb{P}\{|C_t^{(i)} - \ell| < 3/2, \text{ pour tout } i \in [\![1,n]\!]\} \geq 1-\delta.$$

Puisque $\gamma_A + \gamma_B = 1$ et $\ell = (\gamma_A - \gamma_B)m$, nous avons

$$\begin{aligned} |C_t^{(i)} - \ell| &= |C_t^{(i)} - (\gamma_A - \gamma_B)m| = |C_t^{(i)} - (2\gamma_A - 1)m| = |m + C_t^{(i)} - 2m\gamma_A| \\ &= 2m|\omega_A(C_t^{(i)}) - \gamma_A|, \end{aligned} \tag{5.22}$$

donc toujours de manière équivalente

$$\mathbb{P}\{|\omega_A(C_t^{(i)}) - \gamma_A| < 3/(4m), \text{ pour tout } i \in [\![1,n]\!]\} \geq 1-\delta,$$

par conséquent

$$\mathbb{P}\{|\omega_A(C_t^{(i)}) - \gamma_A| < \varepsilon, \text{ pour tout } i \in [\![1,n]\!]\} \geq 1-\delta,$$

ce qu'il fallait démontrer. ■

A partir du théorème 5.2.13, le temps de convergence pour obtenir la proportion γ_A des agents qui étaient dans l'état initial A, avec la précision ε et une probabilité supérieure à $1-\delta$, les deux paramètres ε et δ étant donnés à l'avance, est $O\left(n \log[n/(\varepsilon\delta)]\right)$, donc le temps parallèle de convergence correspondant est $O\left(\log[n/(\varepsilon\delta)]\right)$.

Toujours à partir du théorème 5.2.13, le nombre d'états nécessaires au calcul de la proportion γ_A est égal à $2\lceil 3/(4\varepsilon)\rceil + 1$. Il est important de noter que ce nombre d'états ne dépend pas de n.

Nous allons maintenant montrer que, lorsque le protocole de proportion converge à un instant, alors il converge aussi à tous les instants suivants.

Lemme 5.2.14 *Soient $\varepsilon \in]0,1[$ et l'événement E_t défini par*

$$E_t = \left\{ \left| \omega_A\left(C_t^{(i)}\right) - \gamma_A \right| < \varepsilon \text{ pour tout } i \in [\![1,n]\!] \right\},$$

alors la suite (E_t) est croissante.

Preuve. À partir de l'égalité (5.22), on obtient

$$\max_{1\le i\le n}\left(\left|\omega_A\left(C_t^{(i)}\right) - \gamma_A\right|\right) = \frac{\|C_t - L\|_\infty}{2m},$$

donc, $E_t = \{\|C_t - L\|_\infty < 2m\varepsilon\}$ et d'après le lemme 5.2.7 $(\|C_t - L\|_\infty)_{t\ge 0}$ est décroissante.

Par conséquent, la suite (E_t) est croissante, ce qu'il fallait démontrer. ■

5.2.4 Le problème du comptage

Nous allons maintenant montrer que ce qui précède permet d'obtenir de meilleurs résultats en ce qui concerne le problème de comptage introduit dans [54]. Ce problème vise, pour chaque agent, à calculer le nombre exact d'agents ayant commencé dans l'état initial A. En utilisant les règles d'interaction données par la relation (5.15) et la fonction de sortie

$$\omega'_A(x) = \lfloor n(m+x)/(2m) + 1/2 \rfloor,$$

nous pouvons exploiter les résultats obtenus dans la section précédente pour montrer que le problème de comptage peut être résolu avec $O(n)$ états, avec un temps parallèle toujours égal à $O(\log n)$. Le protocole de comptage est donc défini par $(\Sigma, Q, \Xi', \iota, f, \omega'_A)$. Les ensembles Σ,Q et les fonctions ι,f sont définis de la même façon que dans le protocole de proportion. ω'_A a déjà été définie, l'ensemble de sortie est $\Xi' = [\![0,n]\!]$.

Théorème 5.2.15 *Pour tout $\delta \in]0;1[$ et $t \ge n\left(5.12 \ln n - 6.59 \ln \delta + 2.58\right)$, nous avons, en prenant $m = \lceil 3n/2 \rceil$,*

$$\mathbb{P}\left\{\omega'_A(C_t^{(i)}) = n_A, \forall i \in [\![1,n]\!]\right\} \ge 1-\delta.$$

Preuve. Observons que

$$\omega'_A(x) = \lfloor n\omega_A(x) + 1/2 \rfloor.$$

En appliquant le théorème 5.2.13 avec $\varepsilon = 1/(2n)$ et pour $t \geq n\,(5.12 \ln n - 6.59 \ln \delta + 2.58)$, nous obtenons

$$\mathbb{P}\{|n\omega_A(C_t^{(i)}) - n\gamma_A| < 1/2 \text{ pour tout } i = 1, \ldots, n\} \geq 1 - \delta.$$

Comme $n\gamma_A = n_A$ est un entier, nous avons

$$\mathbb{P}\{\omega'_A(C_t^{(i)}) = n_A, \text{ pour tout } i = 1, \ldots, n\} \geq 1 - \delta,$$

ce qu'il fallait démontrer. ■

Ainsi, chaque agent résout le problème du comptage avec $O(n)$ états en un temps parallèle de $O(\log n)$ avec probabilité élevée.

5.2.5 Caractère optimal de notre protocole

Nous allons montrer que $\Omega(n)$ états et $\Omega(\log n)$ en temps parallèle sont des bornes inférieures pour le problème de comptage, ce qui indique que notre protocole est optimal pour résoudre le problème du comptage. Enfin, nous prouvons que tout algorithme résolvant le problème de proportion avec une précision $\varepsilon \in]0; 1[$ nécessite $\Omega(1/\varepsilon)$ états. Cela démontre que notre protocole de proportion est optimal du point de vue du nombre d'états.

Théorème 5.2.16 *Tout algorithme résolvant le problème de comptage nécessite un temps parallèle de* $\Omega(\log n)$ *pour atteindre la convergence.*

Preuve. Résoudre le problème du comptage permet de résoudre le problème de la majorité exacte. Or, d'après le théorème C.1 de [8], résoudre la majorité nécessite $\Omega(\log n)$ en temps parallèle pour atteindre la convergence dans le pire des cas, par conséquent le comptage nécessite bien $\Omega(\log n)$ en temps parallèle pour atteindre la convergence. ■

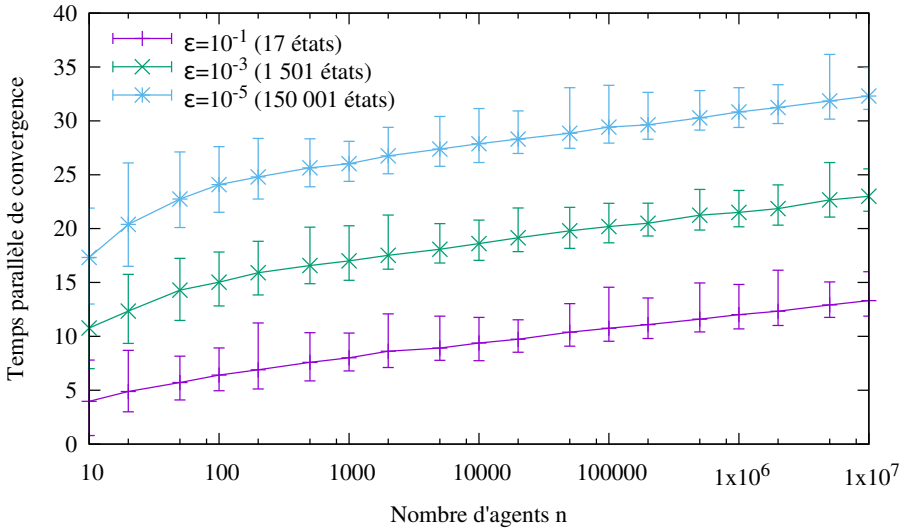
Théorème 5.2.17 *Tout algorithme résolvant le problème de comptage nécessite* $\Omega(n)$ *états.*

Preuve. Pour résoudre le problème de comptage, la taille de l'ensemble de sortie doit être $n + 1$, afin de pouvoir différencier tous les cas. Aussi le nombre d'états (c'est-à-dire $|Q|$) doit être d'au moins $n + 1$. Par conséquent la borne inférieure du nombre d'états est $\Omega(n)$. ■

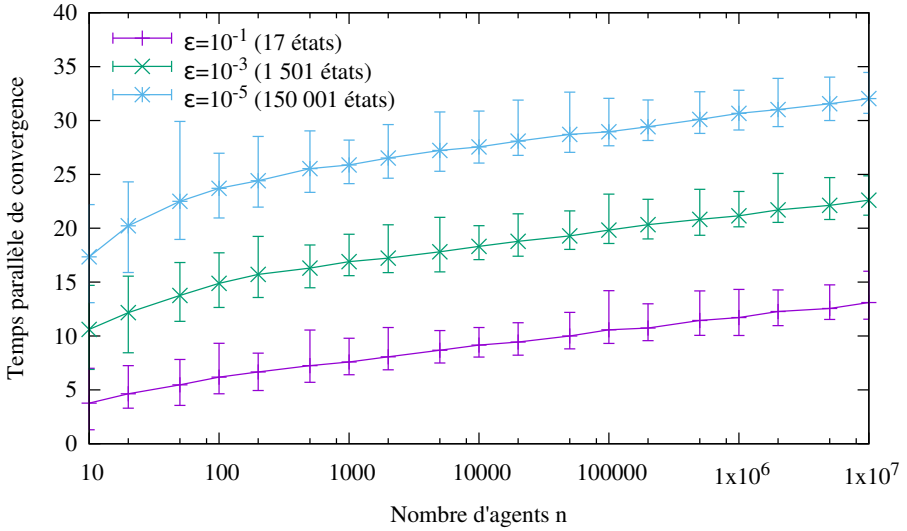
Théorème 5.2.18 *Tout algorithme résolvant le problème de la proportion avec une précision de* $\varepsilon \in]0; 1[$ *nécessite* $\Omega(1/\varepsilon)$ *états.*

Preuve. La valeur de γ_A peut être n'importe quelle valeur rationnelle entre 0 et 1, la différence entre deux valeurs possibles de sortie ne peut excéder 2ε, par conséquent une borne inférieure pour la taille de l'ensemble de sortie Ξ est $\lceil 1/(2\varepsilon) \rceil + 1$. Donc, le nombre d'états de travail (c'est-à-dire $|Q|$) est au moins $\lceil 1/(2\varepsilon) \rceil + 1$. La borne inférieure du nombre d'états est donc $\Omega(1/\varepsilon)$. ■

5.2.6 Résultats de simulations



(a) $\gamma_A = \gamma_B = 1/2$



(b) $\gamma_A = 3/4$ et $\gamma_B = 1/4$

FIGURE 5.3 : Nombre d'interactions par agent pour atteindre la convergence en fonction de la taille du système.

Nous avons effectué des simulations pour illustrer notre analyse théorique. La figure 5.3 fournit un aperçu de ces simulations. Dans ces graphiques, chaque point de la courbe représente la moyenne de 100 simulations (avec le maximum et le minimum de ces 100 simulations). Une simulation consiste à lancer un protocole de population basé sur la moyenne, ce protocole s'arrêtant à la convergence, c'est-à-dire quand la distance maximale entre les valeurs de deux agents est inférieure ou égale à 2, ou, ce qui revient au même, quand tous les agents ont convergé vers γ_A avec une précision de ε. Le nombre n d'agents varie de 10 à 10^7, et la précision du résultat peut prendre les valeurs 10^{-1}, 10^{-3}, et 10^{-5}. Comme cela a été montré théoriquement, les

figures 5.3(a) et 5.3(b) illustrent le fait que le nombre d'interactions par agent pour converger est indépendant de la valeur de γ_A ou γ_B, c'est-à-dire qu'il est indépendant de la proportion initiale de A ou de B.

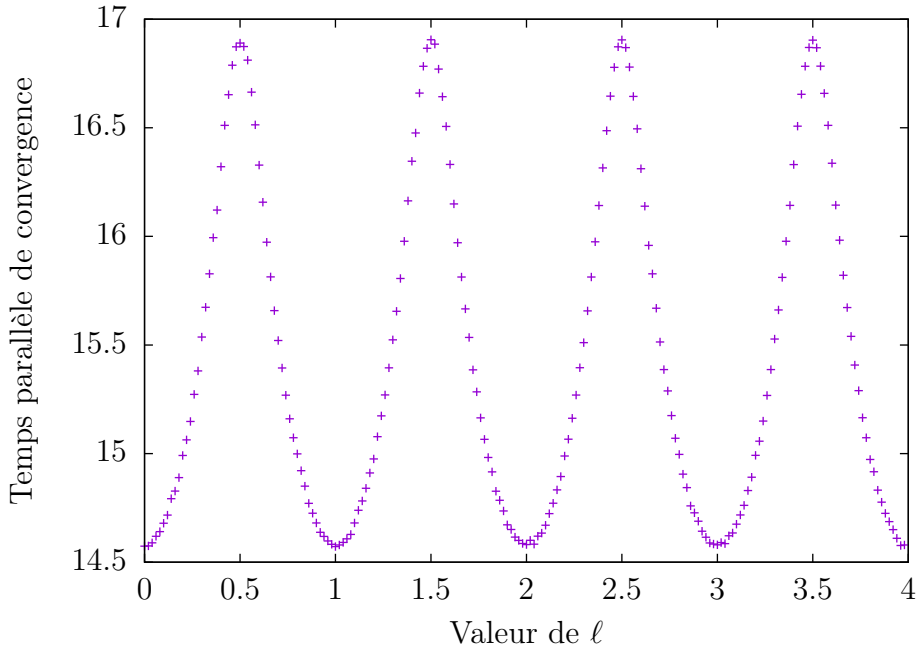


FIGURE 5.4 : Temps de convergence en fonction de ℓ pour $n = 10^4$

Le théorème 5.2.8, traitant du cas particulier où $\ell - \lfloor \ell \rfloor = 1/2$, est intéressant dans la mesure où avec une démonstration simple, on obtient des coefficients nettement inférieurs au cas général. Nous avons mis en évidence, au moyen de nombreuses simulations, l'influence de la partie fractionnaire de ℓ sur le temps de convergence. Nous avons donc effectué pour différentes répartitions initiales, 10^4 simulations dans un système à 10^4 noeuds et 201 états c'est-à-dire que $m = 100$. Nous avons fait varier le nombre de noeuds dans l'état A de 5000 à 5200, le nombre de noeuds dans l'état B variant dans le même temps de 5000 à 4800, en conséquence de quoi ℓ varie de 0 à 4 avec un pas de 0.02. Le résultat est illustré par la figure 5.4. Il apparaît clairement que le cas où $\ell - \lfloor \ell \rfloor = 1/2$ est le plus défavorable, c'est-à-dire que le temps de convergence est plus long quand la partie fractionnaire de ℓ s'approche de 0.5, par conséquent nous pouvons conjecturer que la formule obtenue au théorème 5.2.8 constitue une borne supérieure du temps de convergence du protocole basé sur la moyenne.

5.2.7 Conclusion

Cette section a montré que, dans un système à grande échelle, n'importe quel agent peut calculer rapidement et avec une grande précision spécifiée à l'avance, la proportion d'agents qui ont démarré initialement dans un état d'entrée donné. Ce problème est une généralisation du problème de la majorité. En particulier, notre protocole garantit qu'en utilisant $2\lceil 3/(4\varepsilon) \rceil + 1$ états, tout agent est capable

de calculer la proportion avec une précision de $\varepsilon \in]0;1[$, en utilisant moins de $(3.12 \ln n - 2 \ln \varepsilon - 6.59 \ln \delta + 1.88)$ interactions et avec une probabilité d'au moins $1 - \delta$, pour tout $\delta \in]0;1[$. Nous avons également montré que notre solution est optimale à la fois en temps et en espace.

Chapitre 6

Horloge globale

Dans ce chapitre nous cherchons à construire une horloge globale dans le cadre des protocoles de population, telle que décrite en section 3.5. L'état de chaque agents consiste en un compteur initialisé à 0 au début, la valeur de ce compteur devant représenté le temps parallèle écoulé depuis l'origine. Nous parlons essentiellement du paradigme à deux choix. Appliqué aux protocoles de population, ce paradigme consiste, à chaque interaction, à incrémenter le compteur le moins élevé, le paradigme à un choix consiste à incrémenter le compteur de l'agent à initiative de l'interaction.

6.1 Introduction

Dans ce chapitre, nous abordons l'importante question de l'analyse du paradigme à deux choix (two-choice paradigm) [53]. Pour illustrer le paradigme multi-choix, supposons que nous ayons un ensemble de m boules qui doivent être placées séquentiellement dans n urnes. Le paradigme "d-choice" pour $d \geq 1$, consiste, pour chaque boule, à choisir aléatoirement, selon une loi uniforme, d urnes parmi les n et, à placer la boule dans la moins remplie des d urnes choisies. Azar et al. [16] ont caractérisé ce problème par ces trois valeurs (m, n, d). Deux questions peuvent se poser, d'une part l'analyse du nombre de boules dans l'urne la plus remplie (charge maximale), et d'autre part, l'écart maximal qui peut exister entre l'urne la plus remplie et la moins remplie (gap). Il a été prouvé que dans le cas le plus simple où $d = 1$ (voir par exemple [70]), la charge maximale est égale à $m/n + \Theta\left(\sqrt{(m/n)\ln n}\right)$, conduisant à un écart qui augmente avec la racine carrée de m. Si, au lieu de choisir une seule urne au hasard, d urnes ($d \geq 2$) sont choisies indépendamment et aléatoirement, et l'urne la moins chargée parmi les d urnes reçoit la boule, alors Azar et al. [16] ont montré que lorsque $m = n$ la charge maximale est $\ln(\ln(n))/\ln(2) + O(1)$ et le plus grand écart est également égal à $\ln(\ln(n))/\ln(2)+O(1)$. Ces résultats montrent qu'en introduisant simplement un choix minimal, nous obtenons une charge équilibrée, ce qui améliore considérablement le résultat. Nous pouvons citer Mitzenmacher et al. [53], "avoir seulement deux choix aléatoires (c'est-à-dire $d = 2$) donne une réduction importante de la charge maximale par rapport à un seul choix, tandis que chaque choix supplémentaire à deux diminue la charge maximale de seulement un facteur constant". D'où le nom du paradigme à deux choix. Plus tard, Berenbrink et al. [18]

ont étudié le cas (m, n, d) pour $d \geq 2$ et $m \gg n$, et prouvé que la charge maximale est égale à $m/n + O(\ln(\ln(n)))$. Notons qu'une preuve plus simple de ce résultat a été récemment trouvée par Talwar et Wieder [74]. Très récemment, Peres et al., [66, 67], en utilisant une mesure basée sur le cosinus hyperbolique, ont généralisé ce problème avec le choix $(1+\beta)$ ($(1+\beta)$-choice). Le choix $(1+\beta)$ consiste, avec une probabilité $1-\beta$, à choisir une seule urne selon une loi uniforme et à y placer la boule (dans ce cas $d = 1$), et, avec une probabilité β, à choisir deux urnes toujours selon une loi uniforme et à placer la boule dans l'urne la moins pleine des deux (dans ce cas $d = 2$). Le nom vient du fait que $\mathbb{E}(d) = 1+\beta$. On peut noter que dans leur modèle, chaque boule possède un poids aléatoire. Ils ont montré qu'il existe une borne logarithmique en n à la fois pour l'écart entre l'urne la plus chargée et la moyenne des charges [66], et pour l'écart entre l'urne la plus chargée et la moins chargée [67]. Dans les deux cas, l'écart est de $O(\log(n)/\beta)$.

Le paradigme à deux choix peut être utilisé dans une multitude d'applications, y compris l'allocation de ressources en ligne équilibrée (où les tâches doivent être allouées de manière dynamique au processeur le moins chargé) [2, 17, 10], l'équilibrage de charge [52, 3, 18] et, très récemment, les protocoles de population [6]. Dans ce dernier cas, le modèle régissant l'évolution de ces systèmes consiste à remplacer chaque urne par un agent et à remplacer le nombre de boules dans l'urne par un compteur. Ainsi, le choix des deux urnes devient le choix des agents interagissant. Au lieu de placer la boule dans l'urne la moins chargée, on incrémente le compteur des agents interagissant le moins élevé. La moyenne de la valeur des compteurs représente le temps parallèle. Pour connaître la précision de l'évaluation du temps parallèle par chaque agent, nous cherchons à borner l'écart maximal entre deux compteurs quelconques, nous appelons cet écart le gap. À l'instant t, ce gap est notée $\mathrm{Gap}(t)$. De nombreux travaux ont été consacrés à la recherche d'approximations asymptotiques du gap pour les grandes valeurs de n. Dans ce chapitre, nous allons plus loin en montrant que pour tout $t \geq 0$, $n \geq 2$ et $\sigma > 0$,

$$\mathbb{P}\left\{\mathrm{Gap}(t) \geq a(1+\sigma)\ln(n) + b\right\} \leq \frac{1}{n^{\sigma}}, \tag{6.1}$$

où les constantes a et b, qui sont indépendantes de t, σ et n, sont optimisées et données explicitement, ce qui, à notre connaissance, n'avait jamais été fait auparavant.

Le reste de ce chapitre est structuré comme suit. Dans la section 6.2, nous présentons le problème qui nous intéresse et un algorithme simple pour le résoudre. La section 6.3 est la principale contribution de notre travail qui consiste à fournir une limite explicite de la distribution de l'écart entre deux nœuds quelconques. La section 6.4 évalue les constantes a et b obtenues par notre analyse et les compare aux constantes dérivées du travail de [10]. Le gain de précision obtenu par notre analyse est significatif. Enfin, la section 6.5 fournit un résumé des résultats des expérimentations.

6.2 Description du problème

Chaque agent possède un compteur local, initialisé à 0. Les agents communiquent via des interactions par paires aléatoires. À chaque interaction, les deux agents en inter-

action comparent leurs compteurs, et celui ayant la valeur de compteur la plus faible incrémente son compteur local. L'objectif de cet algorithme simple est la construction d'une horloge globale en garantissant que les valeurs de tous les compteurs d'agents sont concentrées, avec un écart (gap) entre les valeurs minimales et maximales borné par la relation (6.1). Comme les interactions sont aléatoires selon une loi uniforme, ce protocole est équivalent au processus d'équilibrage de charge classique à deux choix [67, 6]. Le but de ce chapitre est de borner le gap, en évaluant avec précision les constantes a et b. Dans la mesure où l'on connaît avec précision l'écart maximal entre deux compteurs, d'autres protocoles de population peuvent utiliser ce compteur comme une *horloge globale* pour effectuer des actions de manière synchronisée et probabiliste, c'est d'ailleurs ce que nous ferons au chapitre 7.

Nous notons $C_t^{(i)}$ l'état de l'agent i à l'instant t. Nous introduisons le processus stochastique $C = \{C_t,\ t \geq 0\}$ où $C_t = \left(C_t^{(1)}, \ldots, C_t^{(n)}\right)$ représente le vecteur d'état du système à l'instant t.

Le choix de deux agents interagissant est fait en utilisant une probabilité uniforme. Étant donné un couple (i, j) d'agents interagissant à l'instant t, les états des agents évoluent de la façon suivante

$$\left(C_{t+1}^{(i)}, C_{t+1}^{(j)}\right) = \begin{cases} \left(C_t^{(i)} + 1, C_t^{(j)}\right) & \text{if } C_t^{(i)} \leq C_t^{(j)} \\ \\ \left(C_t^{(i)}, C_t^{(j)} + 1\right) & \text{if } C_t^{(i)} \geq C_t^{(j)}. \end{cases}$$

Notons que dans le cas où les agents i et j interagissent à l'instant t avec $C_t^{(i)} = C_t^{(j)}$ n'importe lequel des deux agents peut être choisi pour avoir sa valeur incrémentée de 1 à l'instant $t + 1$. Un choix particulier sera présenté par la suite.

L'ensemble des états du processus C est donc $\mathbb{N}^n$ et l'état de ce processus stochastique est aussi appelé configuration du protocole. A l'instant $t = 0$, nous avons $C_0^{(i)} = 0$, pour tout $i \in [\![1, n]\!]$. A chaque instant la valeur d'un seul agent est incrémenté de 1 ce qui signifie que nous avons pour tout $t \geq 0$,

$$\sum_{i=1}^{n} C_t^{(i)} = t.$$

Pour tout $i \in [\![1, n]\!]$, nous introduisons la valeur $x_i(t) = C_t^{(i)} - t/n$, ce qui mène, pour tout $t \geq 0$, à

$$\sum_{i=1}^{n} x_i(t) = 0.$$

La valeur $C_t^{(i)}$ maintenue par l'agent i correspond à sa propre vue de l'horloge globale t du système divisé par n. Plus précisément, une approximation du temps t, fourni par l'agent i, est $nC_t^{(i)}$.

A chaque instant discret $t \geq 0$, deux agents d'indices i et j sont choisis selon une loi uniforme pour interagir, indépendamment du vecteur d'état avec la probabilité $1/(n(n-1))$.

Afin de simplifier la présentation, nous supposons sans perte de généralité qu'à chaque instant t, les valeurs de $x_i(t)$ sont réordonnées de manière décroissante, en

assignant un ordre arbitraire aux agents ayant la même valeur. Plus précisément, à l'instant t, le réordonnancement donne

$$x_1(t) = \max_{i=1,\dots,n} (C_t^{(i)} - t/n) \geq \cdots \geq x_n(t) = \min_{i=1,\dots,n} (C_t^{(i)} - t/n).$$

Nous notons X le rang de l'agent dont la valeur est incrémentée quand une interaction a lieu entre deux agents. Dans le cas d'une interaction de deux agents d'indices i et j sur le vecteur C_t, tel que $C_t^{(i)} = C_t^{(j)}$, nous choisissons d'incrémenter de 1 celui qui a le rang le plus élevé sur le vecteur $x(t)$. Si X_1 et X_2 sont les rangs (sur le vecteur $x(t)$) des agents qui interagissent, alors la probabilité p_ℓ qu'un agent de rang ℓ (sur le vecteur $x(t)$) soit incrémenté, est donnée, pour $\ell \in [\![1, n]\!]$, par

$$\begin{aligned} p_\ell &= \mathbb{P}\{X = \ell\} = \mathbb{P}\{X_1 = \ell, X_2 < \ell\} + \mathbb{P}\{X_1 < \ell, X_2 = \ell\} \\ &= \frac{1}{n}\left(\frac{\ell-1}{n-1}\right) + \left(\frac{\ell-1}{n}\right)\frac{1}{n-1} = \frac{2(\ell-1)}{n(n-1)}. \end{aligned} \tag{6.2}$$

Comme mentionné dans l'introduction, l'objectif du chapitre est l'évaluation de la distribution de la différence entre le maximum et le minimum des entrées du vecteur C_t. Cette différence, notée $\mathrm{Gap}(t)$, est donnée, pour $t \in \mathbb{N}$, par

$$\mathrm{Gap}(t) = \max_{1\leq i\leq n} C_t^{(i)} - \min_{1\leq i\leq n} C_t^{(i)} = x_1(t) - x_n(t).$$

Afin de borner la fonction de répartition de $\mathrm{Gap}(t)$, nous introduisons les fonctions de potentiel suivantes, définies, pour $\alpha \in \mathbb{R}^*$, par

$$\Phi(t) = \sum_{i=1}^{n} e^{\alpha x_i(t)},\ \Psi(t) = \sum_{i=1}^{n} e^{-\alpha x_i(t)} \text{ et } \Gamma(t) = \Phi(t) + \Psi(t).$$

L'utilisation de ces deux fonctions a été proposée par Y. Peres et al. dans [67]. Afin de simplifier les expressions nous définissons aussi

$$\Delta\Phi(t) = \Phi(t+1) - \Phi(t),\ \Delta\Psi(t) = \Psi(t+1) - \Psi(t),\ \Delta\Gamma(t) = \Gamma(t+1) - \Gamma(t).$$

La fonction de potentiel $\Gamma(t)$ est en relation avec la fonction $\mathrm{Gap}(t)$ par le lemme suivant.

Lemme 6.2.1 *Pour tout $t \geq 0$, nous avons*

$$\Gamma(t) \geq 2e^{\alpha Gap(t)/2}. \tag{6.3}$$

Preuve. La fonction exponentielle étant convexe, nous avons, pour tout $a, b \in \mathbb{R}$, $2e^{(a+b)/2} \leq e^a + e^b$. Puisque $\mathrm{Gap}(t) = x_1(t) - x_n(t)$, nous obtenons

$$\Gamma(t) = \sum_{i=1}^{n} e^{\alpha x_i(t)} + \sum_{i=1}^{n} e^{-\alpha x_i(t)} \geq e^{\alpha x_1(t)} + e^{-\alpha x_n(t)} \geq 2e^{\alpha(x_1(t)-x_n(t))/2} = 2e^{\alpha \mathrm{Gap}(t)/2},$$

ce qu'il fallait démontrer. ■

Ce résultat sera utilisé à la fin du chapitre pour l'évaluation de la distribution de $\mathrm{Gap}(t)$ qui est basée sur l'évaluation de celle de $\Gamma(t)$, cette dernière formant la partie principale du chapitre.

6.3 Analyse

Nous avons besoin de deux lemmes techniques qui sont démontrés dans l'annexe C.

Lemme 6.3.1 *Pour tout $x \in \mathbb{R}$, nous avons $1 + x \leq e^x$. Pour tout $x \in]-\infty, c]$, nous avons $e^x \leq 1 + x + x^2$, où c est l'unique solution non nulle de l'équation $e^c - 1 - c - c^2 = 0$. La valeur de c vérifie $1.79 < c < 1.8$.*

Preuve. Voir annexe section C. ■

Lemme 6.3.2 *Soient $u = (u_k)_{k\geq 1}$ et $v = (v_k)_{k\geq 1}$, deux suites monotones de nombres réels et soit $(m_n)_{n\geq 1}$, la suite des valeurs moyennes de la suite v définie, pour $n \geq 1$, par*

$$m_n = \frac{1}{n}\sum_{k=1}^{n} v_k.$$

Si les deux suites u et v sont toutes les deux croissantes ou toutes les deux décroissantes, nous avons

$$\sum_{k=1}^{n} u_k v_k \geq m_n \sum_{k=1}^{n} u_k.$$

Si une des deux suites est croissante et l'autre est décroissante alors nous avons

$$\sum_{k=1}^{n} u_k v_k \leq m_n \sum_{k=1}^{n} u_k.$$

Preuve. Voir annexe section C. ■

Pour tout $t \geq 0$, nous introduisons la notation $x(t) = (x_1(t), \ldots, x_n(t))$.

Lemme 6.3.3 *Pour tout $\alpha \in]-1; 1[$, nous avons*

$$\mathbb{E}(\Delta\Phi(t) \mid x(t)) \leq \left(\alpha + \alpha^2\left(1 - \frac{2}{n}\right)\right)\sum_{i=1}^{n} p_i e^{\alpha x_i} - \left(\frac{\alpha}{n} - \frac{\alpha^2}{n^2}\right)\Phi(t). \tag{6.4}$$

Preuve. Comme les $x_i(t)$ sont ordonnés, ils peuvent changer leur valeur et leur ordre à chaque instant. Par conséquent, nous pouvons définir une permutation sur $[\![1, n]\!]$ noté σ_t telle que, pour tout $u \in [\![1, n]\!]$, si $x_i(t) = C_t^{(u)} - t/n$ alors $x_{\sigma_t(i)}(t+1) = C_{t+1}^{(u)} - (t+1)/n$. Supposons que le rang de l'agent u, dont la valeur est incrémentée à l'instant t, est égal à i. Dans ce cas nous avons

$$x_{\sigma_t(i)}(t+1) = C_{t+1}^{(u)} - \frac{t+1}{n} = C_t^{(u)} + 1 - \frac{t+1}{n} = C_t^{(u)} - \frac{t}{n} + 1 + \frac{t}{n} - \frac{t+1}{n} = x_i(t) + 1 - \frac{1}{n}.$$

Cela mène, pour tout $i \in [\![1,n]\!]$, à $x_{\sigma_t(i)}(t+1) = x_i(t) + 1_{\{X=i\}} - 1/n$, où 1_A est la fonction indicatrice de l'événement A. Nous obtenons

$$\begin{aligned}\Delta\Phi(t) &= \sum_{i=1}^{n} \left(e^{\alpha x_i(t+1)} - e^{\alpha x_i(t)}\right) \\ &= \sum_{i=1}^{n} \left(e^{\alpha x_{\sigma_t(i)}(t+1)} - e^{\alpha x_i(t)}\right) \\ &= \sum_{i=1}^{n} \left(e^{\alpha(1_{\{X=i\}}-1/n)} - 1\right) e^{\alpha x_i(t)}\end{aligned}$$

En utilisant le fait que $e^x \leq 1 + x + x^2$ pour $x \leq 1$, voir lemme 6.3.1, nous obtenons, du fait que $\alpha(1_{\{X=i\}} - 1/n) \leq 1$,

$$\begin{aligned}e^{\alpha(1_{\{X=i\}}-1/n)} - 1 &\leq \alpha(1_{\{X=i\}} - 1/n) + \alpha^2(1_{\{X=i\}} - 1/n)^2 \\ &= \alpha(1_{\{X=i\}} - 1/n) + \alpha^2\left(1_{\{X=i\}}(1 - \frac{2}{n}) + \frac{1}{n^2}\right) \\ &= \left(\alpha + \alpha^2\left(1 - \frac{2}{n}\right)\right) 1_{\{X=i\}} - \left(\frac{\alpha}{n} - \frac{\alpha^2}{n^2}\right).\end{aligned}$$

En prenant l'espérance de $\Delta\Phi(t)$, connaissant $x(t)$, nous obtenons, sachant que $\mathbb{E}(1_{\{X=i\}}) = p_i$,

$$\begin{aligned}\mathbb{E}(\Delta\Phi(t) \mid x(t)) &\leq \sum_{i=1}^{n} \left[p_i\left(\alpha + \alpha^2\left(1 - \frac{2}{n}\right)\right) - \left(\frac{\alpha}{n} - \frac{\alpha^2}{n^2}\right)\right] e^{\alpha x_i} \\ &= \left(\alpha + \alpha^2\left(1 - \frac{2}{n}\right)\right) \sum_{i=1}^{n} p_i e^{\alpha x_i} - \left(\frac{\alpha}{n} - \frac{\alpha^2}{n^2}\right) \Phi(t),\end{aligned}$$

ce qui termine la preuve. ■

Les relations suivantes seront souvent utilisées par la suite. Puisque, pour $i \in [\![1,n]\!]$, $p_i = 2(i-1)/(n(n-1))$, alors nous avons pour tout $\lambda \in]0,1[$ avec $\lambda n \in \mathbb{N}$,

$$\frac{1}{n}\sum_{i=1}^{n} p_i = \frac{1}{n}, \tag{6.5}$$

$$\frac{1}{\lambda n}\sum_{i=1}^{\lambda n} p_i = \frac{\lambda n - 1}{n(n-1)} \leq \frac{\lambda}{n}, \tag{6.6}$$

$$\frac{1}{(1-\lambda)n}\sum_{i=\lambda n+1}^{n} p_i = \frac{(1+\lambda)n - 1}{n(n-1)} \geq \frac{1+\lambda}{n}. \tag{6.7}$$

Corollaire 6.3.4 *Pour tout $\alpha \in]0,1[$, nous avons*

$$\mathbb{E}(\Delta\Phi(t) \mid x(t)) \leq \frac{\alpha^2}{n}\left(1 - \frac{1}{n}\right)\Phi(t).$$

Preuve. Pour prouver ce résultat, observons que $(e^{\alpha x_i})_i$ est une suite décroissante et que $(p_i)_i$ est une suite croissante. En utilisant la relation (6.5) et en appliquant le lemme 6.3.2 nous obtenons

$$\sum_{i=1}^{n} p_i e^{\alpha x_i(t)} \leq \frac{1}{n}\left(\sum_{i=1}^{n} p_i\right)\left(\sum_{i=1}^{n} e^{\alpha x_i(t)}\right) = \frac{\Phi(t)}{n}.$$

En insérant ce résultat dans l'inégalité (6.4), nous obtenons

$$\begin{aligned}
\mathbb{E}(\Delta\Phi(t) \mid x(t)) &\leq \left(\alpha + \alpha^2\left(1 - \frac{2}{n}\right)\right)\sum_{i=1}^{n} p_i e^{\alpha x_i} - \left(\frac{\alpha}{n} - \frac{\alpha^2}{n^2}\right)\Phi(t) \\
&\leq \left[\frac{\alpha}{n} + \frac{\alpha^2}{n}\left(1 - \frac{2}{n}\right) - \left(\frac{\alpha}{n} - \frac{\alpha^2}{n^2}\right)\right]\Phi(t) \\
&= \frac{\alpha^2}{n}\left(1 - \frac{1}{n}\right)\Phi(t),
\end{aligned}$$

ce qu'il fallait démontrer. ■

Lemme 6.3.5 *Pour tout $\alpha \in]-1, 1[$, nous avons*

$$\mathbb{E}(\Delta\Psi(t) \mid x(t)) \leq \left(-\alpha + \alpha^2\left(1 - \frac{2}{n}\right)\right)\sum_{i=1}^{n} p_i e^{-\alpha x_i} + \left(\frac{\alpha}{n} + \frac{\alpha^2}{n^2}\right)\Psi(t). \tag{6.8}$$

Preuve. Il suffit de remplacer α par $-\alpha$ dans la preuve du lemme 6.3.3. ■

Corollaire 6.3.6 *Pour tout $\alpha \in]0, 1[$, nous avons*

$$\mathbb{E}\left(\Delta\Psi(t) \mid x(t)\right) \leq \frac{\alpha^2}{n}\left(1 - \frac{1}{n}\right)\Psi(t).$$

Preuve. Observons que pour $\alpha \in [0, 1]$, nous avons $-\alpha + \alpha^2(1 - 2/n) \leq 0$. Il s'ensuit que la suite $((-\alpha + \alpha^2(1 - 2/n))e^{-\alpha x_i})_i$ est décroissante. La suite $(p_i)_i$ est croissante, aussi en utilisant la relation (6.5) et en appliquant le lemme 6.3.2 nous obtenons

$$\begin{aligned}
\left(-\alpha + \alpha^2\left(1 - \frac{2}{n}\right)\right)\sum_{i=1}^{n} p_i e^{-\alpha x_i(t)} &\leq \left(-\frac{\alpha}{n} + \frac{\alpha^2}{n}\left(1 - \frac{2}{n}\right)\right)\sum_{i=1}^{n} e^{-\alpha x_i(t)} \\
&= \left(-\frac{\alpha}{n} + \frac{\alpha^2}{n}\left(1 - \frac{2}{n}\right)\right)\Psi(t).
\end{aligned}$$

En insérant ce résultat dans l'inégalité (6.8), nous obtenons

$$\begin{aligned}
\mathbb{E}(\Delta\Psi(t) \mid x(t)) &\leq \left(-\alpha + \alpha^2\left(1 - \frac{2}{n}\right)\right)\sum_{i=1}^{n} p_i e^{-\alpha x_i} + \left(\frac{\alpha}{n} + \frac{\alpha^2}{n^2}\right)\Psi(t) \\
&\leq \left[-\frac{\alpha}{n} + \frac{\alpha^2}{n}\left(1 - \frac{2}{n}\right) + \left(\frac{\alpha}{n} + \frac{\alpha^2}{n^2}\right)\right]\Psi(t) \\
&= \frac{\alpha^2}{n}\left(1 - \frac{1}{n}\right)\Psi(t),
\end{aligned}$$

ce qui termine la preuve. ■

Les deux lemmes précédents qui donnent une limite à l'augmentation des fonctions $\Phi(t)$ et $\Psi(t)$, seront utilisés pour prouver le théorème 6.3.11. La preuve des résultats suit les idées ingénieuses de l'article novateur [67] dans lequel les auteurs prouvent que $\mathrm{Gap}(t)$ est inférieur à $O(\ln(n))$ avec une probabilité élevée. Dans [10], Alistarh et al. fournissent une preuve plus rigoureuse à partir de laquelle nous avons extrait les constantes associées à ce comportement asymptotique. Ces constantes sont données à la fin de la section 6.4. L'idée originale de ce chapitre est de paramétrer autant que possible les preuves pour obtenir les plus petites valeurs des constantes a et b utilisées dans la relation (6.1) qui est prouvée dans le théorème 6.3.13. Le calcul numérique de ces constantes obtenues dans la section 6.4, montre qu'elles sont remarquablement petites par rapport à celles extraites de [10].

Dans la suite, nous introduisons deux paramètres $\mu, \rho \in]0; 1/2[$ (qui sont fixés à $1/4$ dans [67] et dans [10]). Comme (x_i) est décroissante, nous avons $x_{\rho n} \geq x_{(1-\mu)n}$. Les lemmes 6.3.7 et 6.3.8 traitent des cas avec des conditions équilibrées, c'est-à-dire $x_{\rho n} \geq 0$ d'une part et $0 \geq x_{(1-\mu)n}$ d'autre part. Les cas ayant des conditions déséquilibrées sont les cas complémentaires, c'est-à-dire $x_{\rho n} \geq x_{(1-\mu)n} > 0$ d'une part et $0 > x_{\rho n} \geq x_{(1-\mu)n}$ d'autre part. Ces derniers cas sont traités respectivement dans les lemmes 6.3.9 et 6.3.10. Le théorème 6.3.11 examine systématiquement chaque cas pour obtenir une relation de récurrence pour $\mathbb{E}(\Gamma(t))$. Le théorème 6.3.12 utilise cette récurrence pour borner $\mathbb{E}(\Gamma(t))$. Finalement, le théorème 6.3.13 fournit une borne inférieure précise pour $\Gamma(t)$ avec une probabilité élevée. Son corollaire donne le résultat que l'on cherchait, c'est-à-dire la relation (6.1).

Lemme 6.3.7 *Soit $\alpha, \mu \in]0, 1[$ avec $\mu n \in \mathbb{N}$ et $\mu > \alpha/(1+\alpha)$. Si $x_{(1-\mu)n}(t) \leq 0$, alors nous avons*

$$\begin{aligned}\mathbb{E}(\Phi(t+1) \mid x(t)) &\leq \left(1 - \frac{\alpha}{n}\left[\mu - \alpha(1-\mu) + \frac{\alpha(1-2\mu)}{n}\right]\right)\Phi(t) + \alpha + \alpha^2\left(1 - \frac{2}{n}\right) \\ &\leq \left(1 - \frac{\alpha}{n}\left[\mu - \alpha(1-\mu)\right]\right)\Phi(t) + \alpha + \alpha^2. \qquad (6.9)\end{aligned}$$

Preuve. Si $x_{(1-\mu)n} \leq 0$, nous avons $e^{\alpha x_i(t)} \leq 1$ pour tout $i > (1-\mu)n$, aussi

$$\sum_{i=1}^{n} p_i e^{\alpha x_i(t)} \leq \sum_{i=1}^{(1-\mu)n} p_i e^{\alpha x_i(t)} + \sum_{i=(1-\mu)n+1}^{n} p_i \leq \sum_{i=1}^{(1-\mu)n} p_i e^{\alpha x_i(t)} + 1.$$

Nous utilisons maintenant le lemme 6.3.3. La suite $\left(e^{\alpha x_i(t)}\right)_i$ est décroissante et la suite $(p_i)_i$ est croissante. En utilisant la relation (6.6) et en appliquant le lemme 6.3.2, nous obtenons

$$\sum_{i=1}^{(1-\mu)n} p_i e^{\alpha x_i(t)} \leq \frac{(1-\mu)n - 1}{n(n-1)}\left(\sum_{i=1}^{(1-\mu)n} e^{\alpha x_i(t)}\right) \leq \frac{(1-\mu)\Phi(t)}{n}$$

et aussi

$$\sum_{i=1}^{n} p_i e^{\alpha x_i(t)} \leq \frac{(1-\mu)\Phi(t)}{n} + 1.$$

En insérant cette borne dans l'inégalité 6.3.3, nous obtenons

$$\begin{aligned}\mathbb{E}(\Delta\Phi(t) \mid x(t)) &\leq \left(\alpha+\alpha^2\left(1-\frac{2}{n}\right)\right)\sum_{i=1}^{n} p_i e^{\alpha x_i} - \left(\frac{\alpha}{n}-\frac{\alpha^2}{n^2}\right)\Phi(t)\\ &\leq \left(\alpha+\alpha^2\left(1-\frac{2}{n}\right)\right)\left(\frac{(1-\mu)\Phi(t)}{n}+1\right) - \left(\frac{\alpha}{n}-\frac{\alpha^2}{n^2}\right)\Phi(t)\\ &\leq \left[\left(\alpha+\alpha^2\left(1-\frac{2}{n}\right)\right)\left(\frac{1-\mu}{n}\right) - \frac{\alpha}{n}+\frac{\alpha^2}{n^2}\right]\Phi(t) + \alpha+\alpha^2\left(1-\frac{2}{n}\right)\\ &= -\frac{\alpha}{n}\left[\mu-\alpha(1-\mu)+\frac{\alpha(1-2\mu)}{n}\right]\Phi(t) + \alpha+\alpha^2\left(1-\frac{2}{n}\right).\end{aligned}$$

Nous terminons la preuve en constatant que $\mathbb{E}(\Phi(t) \mid x(t)) = \Phi(t)$. La seconde inégalité est immédiate. ■

Un résultat analogue est obtenu pour $\Psi(t)$ dans le lemme suivant.

Lemme 6.3.8 *Soient $\alpha, \rho \in]0;1[$ avec $\rho n \in \mathbb{N}$ et $\rho > \alpha/(1-\alpha)$. Si $x_{\rho n}(t) \geq 0$, alors nous avons*

$$\begin{aligned}\mathbb{E}\left(\Psi(t+1) \mid x(t)\right) &\leq \left(1-\frac{\alpha}{n}\left[\rho-\alpha(1+\rho)+\frac{\alpha(1+2\rho)}{n}\right]\right)\Psi(t) + \alpha\rho(1+\rho)\\ &\leq \left(1-\frac{\alpha}{n}\left[\rho-\alpha(1+\rho)\right]\right)\Psi(t) + \alpha\rho(1+\rho). \qquad (6.10)\end{aligned}$$

Preuve. Pour $\alpha \in]0,1[$, nous avons $-\alpha+\alpha^2(1-2/n) \leq 0$. Par conséquent, nous avons

$$\left(-\alpha+\alpha^2\left(1-\frac{2}{n}\right)\right)\sum_{i=1}^{n} p_i e^{-\alpha x_i(t)} \leq \left(-\alpha+\alpha^2\left(1-\frac{2}{n}\right)\right)\sum_{i=\rho n+1}^{n} p_i e^{-\alpha x_i(t)}.$$

La suite $\left((-\alpha+\alpha^2(1-2/n))e^{-\alpha x_i(t)}\right)_i$ est décroissante et la suite $(p_i)_i$ est croissante, aussi en utilisant la relation (6.7) et en appliquant le lemme 6.3.2 nous obtenons, du fait que $x_{\rho n} \geq 0$,

$$\begin{aligned}&\left(-\alpha+\alpha^2\left(1-\frac{2}{n}\right)\right)\sum_{i=\rho n+1}^{n} p_i e^{-\alpha x_i(t)}\\ &\qquad\leq \left(-\alpha+\alpha^2\left(1-\frac{2}{n}\right)\right)\frac{(1+\rho)n-1}{n(n-1)}\sum_{i=\rho n+1}^{n} e^{-\alpha x_i(t)}\\ &\qquad\leq \left(-\alpha+\alpha^2\left(1-\frac{2}{n}\right)\right)\frac{1+\rho}{n}\left(\Psi(t)-\sum_{i=1}^{\rho n} e^{-\alpha x_i(t)}\right)\\ &\qquad\leq \left(-\alpha+\alpha^2\left(1-\frac{2}{n}\right)\right)\frac{(1+\rho)(\Psi(t)-\rho n)}{n}.\end{aligned}$$

En insérant cette borne dans l'inégalité du lemme 6.3.5, nous obtenons

$$\begin{aligned}
\mathbb{E}(\Delta\Psi(t) \mid x(t)) &= \mathbb{E}(\Psi(t+1) - \Psi(t) \mid x(t)) \\
&\leq \left(-\alpha + \alpha^2\left(1 - \frac{2}{n}\right)\right) \sum_{i=1}^{n} p_i e^{\alpha x_i} + \left(\frac{\alpha}{n} + \frac{\alpha^2}{n^2}\right) \Psi(t) \\
&\leq \left(-\alpha + \alpha^2\left(1 - \frac{2}{n}\right)\right) \frac{(1+\rho)(\Psi(t) - \rho n)}{n} + \left(\frac{\alpha}{n} + \frac{\alpha^2}{n^2}\right) \Psi(t) \\
&\leq -\frac{\alpha}{n}\left[\rho - \alpha(1+\rho) + \frac{\alpha(1+2\rho)}{n}\right] \Psi(t) + \rho\alpha(1+\rho)\left(1 - \alpha\left(1 - \frac{2}{n}\right)\right) \\
&\leq -\frac{\alpha}{n}\left[\rho - \alpha(1+\rho) + \frac{\alpha(1+2\rho)}{n}\right] \Psi(t) + \rho\alpha(1+\rho).
\end{aligned}$$

Nous terminons la preuve en constatant que $\mathbb{E}(\Psi(t) \mid x(t)) = \Psi(t)$. La seconde inégalité est immédiate. ■

Lemme 6.3.9 *Soient $\alpha, \mu \in]0; 1/2[$ avec $\mu n \in \mathbb{N}$ et $\mu \in]\alpha/(1+\alpha), (1-2\alpha)/(1-\alpha)[$, soit $\mu' \in]0, 1[$ avec $\mu' n \in \mathbb{N}$ et $\mu' \in]\mu/(1-\mu), 1/(1+\alpha)[$ et soit $\gamma_1 \in]0, 1[$.*

Si $x_{(1-\mu)n} > 0$ et $\mathbb{E}(\Delta\Phi(t) \mid x(t)) \geq -\left(1 - \mu'(\alpha+1)\right) \dfrac{\alpha\gamma_1}{n}\Phi(t)$ et $\Phi(t) \geq \lambda_1\Psi(t)$, alors nous avons

$$\Gamma(t) \leq c_1 n,$$

où

$$c_1 = \left(1 + \frac{1}{\lambda_1}\right) C_1 \left(\frac{C_1}{\mu\lambda_1}\right)^{\mu/((1-\mu)\mu' - \mu)}, \; C_1 = \frac{(1-\mu')\,(2+\alpha)}{(1-\gamma_1)\,(1-\mu'(1+\alpha))},$$

$$\text{et } \lambda_1 = \frac{1 - \mu - \alpha(2-\mu)}{2\alpha}.$$

La condition $\mu < (1-2\alpha)/(1-\alpha)$ est nécessaire pour que $\lambda_1 > 0$. La valeur de λ_1 sera utilisée dans le théorème 6.3.11. La condition $\mu' > \mu/(1-\mu)$ est nécessaire pour que la puissance utilisée dans le calcul de c_1 soit positive.

Preuve. Voir annexe section C. ■

Lemme 6.3.10 *Soient $\alpha, \rho \in]0; 1/2[$ avec $\rho n \in \mathbb{N}$ et $\rho \in]\alpha/(1-\alpha); 1/(1+\alpha)[$, soit $\rho' \in]\rho/(1-\rho), (1-2\alpha)/(1-\alpha)[$ avec $\rho' n \in \mathbb{N}$ et soit $\gamma_2 \in]0, 1[$.*

Si $x_{\rho n} < 0$ et $\mathbb{E}(\Delta\Psi(t) \mid x(t)) \geq -\left[1 - 2\alpha - \rho'(1-\alpha)\right] \dfrac{\alpha\gamma_2}{n}\Psi(t)$ et $\Psi(t) \geq \lambda_2\Phi(t)$, alors nous avons

$$\Gamma(t) \leq c_2 n,$$

où

$$c_2 = \left(1 + \frac{1}{\lambda_2}\right) C_2 \left(\frac{C_2}{\rho\lambda_2}\right)^{\rho/((1-\rho)\rho' - \rho)}, \; C_2 = \frac{(1-\rho')\,(2 - 2\alpha - \rho'(1-\alpha))}{(1-\gamma_2)\,(1 - 2\alpha - \rho'(1-\alpha))},$$

$$et\ \lambda_2 = \frac{1-\rho(1+\alpha)}{2\alpha}.$$

La condition $\rho < 1/(1+\alpha)$ est nécessaire pour que $\lambda_2 > 0$. La valeur de λ_2 sera utilisée dans le théorème 6.3.11. La condition $\rho' > \rho/(1-\rho)$ est nécessaire pour que la puissance utilisée dans le calcul de c_2 soit positive.

Preuve. Voir annexe section C. ■

Théorème 6.3.11 *Soient $\alpha, \mu, \rho \in]0, 1/2[$ avec $\mu n, \rho n \in \mathbb{N}$, $\mu \in (\alpha/(1+\alpha), (1-2\alpha)/(1-\alpha))$ et $\rho \in (\alpha/(1-\alpha), 1/(1+\alpha))$. Soit $\mu' \in (\mu/(1-\mu), 1/(1+\alpha))$ avec $\mu' n \in \mathbb{N}$ et soit $\rho' \in (\rho/(1-\rho), (1-2\alpha)/(1-\alpha))$ avec $\rho' n \in \mathbb{N}$. Soient $\gamma_1, \gamma_2 \in (0,1)$. Nous avons alors*

$$\mathbb{E}(\Gamma(t+1) \mid x(t)) \leq \left(1 - c_4 \frac{\alpha}{n}\right) \Gamma(t) + c_3,$$

où

$$c_4 = \min \left\{ \mu - \alpha(1-\mu), \rho - \alpha(1+\rho), \gamma_1 \left(1 - \mu'(\alpha+1)\right), \frac{\alpha\left(1-\mu-\alpha(2-\mu)\right)}{1-\mu(1-\alpha)}, \right.$$
$$\left. \gamma_2 \left(1 - 2\alpha - \rho'(1-\alpha)\right), \frac{\alpha\left(1-\rho(1+\alpha)\right)}{1-\rho(1-\alpha)+2\alpha} \right\}$$

et

$$c_3 = \max\left\{\alpha\left(1+\alpha+\rho(1+\rho)\right), \alpha(1-\mu)(2-\mu), (\alpha+c_4)\alpha c_1, \alpha+\alpha^2, (\alpha+c_4)\alpha c_2\right\},$$

dans lequel

$$c_1 = \left(1 + \frac{1}{\lambda_1}\right) C_1 \left(\frac{C_1}{\mu\lambda_1}\right)^{\mu/((1-\mu)\mu'-\mu)},\ C_1 = \frac{(1-\mu')(2+\alpha)}{(1-\gamma_1)(1-\mu'(1+\alpha))},$$
$$\lambda_1 = \frac{1-\mu-\alpha(2-\mu)}{2\alpha}$$

et

$$c_2 = \left(1 + \frac{1}{\lambda_2}\right) C_2 \left(\frac{C_2}{\rho\lambda_2}\right)^{\rho/((1-\rho)\rho'-\rho)},\ C_2 = \frac{(1-\rho')(2-2\alpha-\rho'(1-\alpha))}{(1-\gamma_2)(1-2\alpha-\rho'(1-\alpha))},$$
$$\lambda_2 = \frac{1-\rho(1+\alpha)}{2\alpha}.$$

Preuve. Voir annexe section C. ■

Nous sommes maintenant capables de proposer une borne supérieure de l'espérance de $\Gamma(t)$.

Théorème 6.3.12 *Pour tout $t \geq 0$, sous les hypothèses du théorème 6.3.11, nous avons*

$$\mathbb{E}(\Gamma(t)) \leq c_3 n/(\alpha c_4)$$

.

Preuve. Nous prouvons ce résultat par récurrence. Pour $t = 0$, nous avons $\Gamma(0) = 2n$. De plus, nous avons

$$c_3 \geq \alpha\left(1 + \alpha + \rho(1 + \rho)\right) \geq \alpha \text{ et } c_4 \leq \mu - \alpha(1 - \mu) \leq \mu \leq 1/2,$$

ce qui implique que $c_3/(\alpha c_4) \geq 2$. Par conséquent, nous avons $\mathbb{E}(\Gamma(0)) = 2n \leq c_3 n/(\alpha c_4)$.

Maintenant, supposons que le résultat soit vrai pour une valeur de $t \geq 0$. A partir du théorème 6.3.11, nous avons

$$\begin{aligned}
\mathbb{E}(\Gamma(t+1)) &= \mathbb{E}\left(\mathbb{E}(\Gamma(t+1) \mid x(t))\right) \\
&\leq \mathbb{E}\left(\left(1 - c_4\frac{\alpha}{n}\right)\Gamma(t) + c_3\right) \\
&\leq \left(1 - c_4\frac{\alpha}{n}\right)\frac{c_3}{\alpha c_4}n + c_3 \\
&= \frac{c_3}{\alpha c_4}n,
\end{aligned}$$

La récurrence est établie donc, pour tout $t \geq 0$,

$$\mathbb{E}(\Gamma(t)) \leq c_3 n/(\alpha c_4),$$

ce qu'il fallait démontrer. ■

Théorème 6.3.13 *Pour tout $t \geq 0$ et $\sigma > 0$, sous les hypothèses du théorème 6.3.11, nous avons*

$$\mathbb{P}\left\{\text{Gap}(t) \geq \frac{2(1+\sigma)}{\alpha}\ln(n) + \frac{2}{\alpha}\ln\left(\frac{c_3}{2\alpha c_4}\right)\right\} \leq \frac{1}{n^\sigma}.$$

Preuve. A partir du lemme 6.2.1 et du théorème 6.3.12, nous avons

$$\Gamma(t) \geq 2e^{\alpha\text{Gap}(t)/2} \text{ et } \frac{c_3 n}{\alpha c_4} \geq \mathbb{E}(\Gamma(t)).$$

Il s'ensuit que

$$2e^{\alpha\text{Gap}(t)/2} \geq n^\sigma\frac{c_3 n}{\alpha c_4} \Longrightarrow \Gamma(t) \geq n^\sigma\frac{c_3 n}{\alpha c_4} \Longrightarrow \Gamma(t) \geq n^\sigma\mathbb{E}(\Gamma(t)).$$

En utilisant l'inégalité de Markov, nous obtenons

$$\begin{aligned}
\mathbb{P}\left\{\text{Gap}(t) \geq \frac{2(\sigma+1)}{\alpha}\ln(n) + \frac{2}{\alpha}\ln\left(\frac{c_3}{2\alpha c_4}\right)\right\} &= \mathbb{P}\left\{2e^{\alpha\text{Gap}(t)/2} \geq n^\sigma\frac{c_3 n}{\alpha c_4}\right\} \\
&\leq \mathbb{P}\left\{\Gamma(t) \geq n^\sigma\mathbb{E}(\Gamma(t))\right\} \leq \frac{1}{n^\sigma},
\end{aligned}$$

ce qu'il fallait démontrer. ■

Le corollaire suivant montre qu'à tout moment, et pour n'importe quel agent, son compteur local se rapproche de l'horloge globale avec une forte probabilité.

Corollaire 6.3.14 *Pour tout $t \geq 0$ et $\sigma > 0$, sous les hypothèses du théorème 6.3.11, nous avons*

$$\mathbb{P}\left\{\left|C_t^{(i)} - \frac{t}{n}\right| < \frac{2(1+\sigma)}{\alpha}\ln(n) + \frac{2}{\alpha}\ln\left(\frac{c_3}{2\alpha c_4}\right),\ \forall i \in [\![1,n]\!]\right\} \geq 1 - \frac{1}{n^\sigma}.$$

Preuve. Par définition, nous avons

$$\forall i \in [\![1,n]\!]\ \exists j \in [\![1,n]\!] \text{ tel que } x_j = C_t^{(i)} - t/n,$$

et comme $x_n \leq 0 \leq x_1$, nous avons $|x_j| \leq x_1 - x_n = \text{Gap}(t)$. Il s'ensuit en utilisant le théorème 6.3.13,

$$\begin{aligned}\mathbb{P}\left\{\left|C_t^{(i)} - \frac{t}{n}\right| \geq \frac{2(1+\sigma)}{\alpha}\ln(n) + \frac{2}{\alpha}\ln\left(\frac{c_3}{2\alpha c_4}\right),\ \forall i \in [\![1,n]\!]\right\} \\ \leq \mathbb{P}\left\{\text{Gap}(t) \geq \frac{2(1+\sigma)}{\alpha}\ln(n) + \frac{2}{\alpha}\ln\left(\frac{c_3}{2\alpha c_4}\right)\right\} \leq \frac{1}{n^\sigma},\end{aligned}$$

ce qu'il fallait démontrer. ■

6.4 Evaluation des constantes

Cette section est dédiée à l'évaluation des constantes a et b de la relation (6.1) et à leur comparaison avec celles déduites de l'analyse de Alistarh et al. [10].
A partir du théorème 6.3.13, nous avons

$$a = \frac{2}{\alpha} \text{ et } b = \frac{2}{\alpha}\ln\left(\frac{c_3}{2\alpha c_4}\right),$$

où c_3 et c_4 sont données par le théorème 6.3.11. Pour commencer, notons que les contraintes du théorème 6.3.11 impliquent l'inégalité suivante : $\rho/(1-\rho) < (1-2\alpha)/(1-\alpha)$, ce qui équivaut à $\rho \leq (1-2\alpha)/(2-3\alpha)$, et qui, combiné avec $\rho \geq \alpha/(1-\alpha)$, mène à $\alpha \leq (5-\sqrt{5})/10 \approx 0.276$.

Pour une valeur fixée de α, nous devons déterminer les valeurs des paramètres $\mu, \rho, \mu', \rho', \gamma_1, \gamma_2$ qui minimisent la constante b. Ceci est réalisé en appliquant un algorithme de Monte-Carlo simple. Le tableau 6.1 montre plusieurs valeurs optimales des constantes a et b, utilisées dans le théorème 6.3.13, et calculées pour plusieurs valeurs de α.

α	0.17	0.18	0.19	0.20	0.21	0.22	0.23	0.24	0.25	0.26	0.27
a	11.77	11.12	10.53	10	9.53	9.10	8.70	8.34	8	7.70	7.41
b	59	63	68	74	82	93	109	134	179	281	739

Tableau 6.1 : Valeurs optimales de a et b en fonction de α

Évaluons maintenant les constantes a et b obtenues dans l'article de Alistarh et al. [10]. Notons que le but de leur travail n'était pas nécessairement axé sur l'optimisation de ces constantes. Néanmoins, comme nous le verrons, l'évaluation

des constantes a et b est une motivation importante de notre travail. A partir des relations (1) et (2) de [10] et comme $\beta = 1$, nous obtenons $0 < \delta \leq \varepsilon = 1/16$ et donc cela donne, pour $\gamma > 0$ et $c \geq 2$,

$$\frac{1+\gamma+c\alpha(1+\gamma)^2}{1-\gamma-c\alpha(1+\gamma)^2} \leq \frac{17}{16},$$

ce qui nous donne,

$$\alpha \leq \frac{1}{33c(1+\gamma)^2} - \frac{1}{c(1+\gamma)^2} \leq \frac{1}{33c(1+\gamma)^2} \leq \frac{1}{66}.$$

En considérant la différence entre les bornes inférieure et supérieure de l'inégalité suivant la relation (11), nous obtenons

$$\exp\left(\frac{\alpha B}{n}\left(3 - \frac{1}{1-\lambda}\right)\right) \leq \frac{16\lambda C(\varepsilon)}{\varepsilon},$$

qui peut aussi s'écrire

$$\exp\left(\frac{\alpha B}{(1-\lambda)n}\right) \leq \left(\frac{16\lambda C(\varepsilon)}{\varepsilon}\right)^{1/(2-3\lambda)}.$$

En utilisant la dernière inégalité obtenue dans la preuve du lemme 4.8, nous obtenons

$$\Gamma(t) \leq \frac{4+\varepsilon}{\varepsilon}\lambda n C(\varepsilon) \exp\left(\frac{\alpha B}{(1-\lambda)n}\right) \leq \frac{4+\varepsilon}{\varepsilon}\lambda n C(\varepsilon)\left(\frac{16\lambda C(\varepsilon)}{\varepsilon}\right)^{1/(2-3\lambda)}.$$

En utilisant ce résultat, à partir du lemme 4.11, nous avons $\mathbb{E}(\Gamma(t)) \leq 4Cn/(\hat{\alpha}\varepsilon)$, où

$$C = \frac{4+\varepsilon}{\varepsilon}\lambda C(\varepsilon)\left(\frac{16\lambda C(\varepsilon)}{\varepsilon}\right)^{1/(2-3\lambda)}, \quad C(\varepsilon) = \frac{(1+\delta)/\lambda - 1 + 3\varepsilon}{3\varepsilon - \varepsilon/3}$$
$$\text{et } \widehat{\alpha} = \alpha(1-\gamma-c\alpha(1+\gamma)^2).$$

En suivant l'idée utilisée pour démontrer le théorème 6.3.13, nous pouvons écrire

$$a = \frac{2}{\alpha} \text{ et } b = \frac{2}{\alpha}\ln\left(\frac{2C}{\widehat{\alpha}\varepsilon}\right).$$

Comme $\alpha \leq 1/66$, nous avons $a \geq 132$. De plus, comme $0 \leq \delta \leq \varepsilon = 1/16$, $\lambda = 2/3 - 1/54 = 35/54$, $\gamma > 0$ et $c \geq 2$, nous obtenons

$$C(\varepsilon) = \frac{(1+\delta)/\lambda - 1 + 3\varepsilon}{3\varepsilon - \varepsilon/3} \geq \frac{1/\lambda - 1 + 3\varepsilon}{3\varepsilon - \varepsilon/3} = \frac{1227}{280}$$

ce qui mène à

$$C = \frac{4+\varepsilon}{\varepsilon}\lambda C(\varepsilon)\left(\frac{16\lambda C(\varepsilon)}{\varepsilon}\right)^{1/(2-3\lambda)} \geq \frac{26585}{144}\left(\frac{6544}{9}\right)^{18}.$$

En ce qui concerne $\widehat{\alpha}$, nous avons $\widehat{\alpha} = \alpha(1-\gamma-c\alpha(1+\gamma)^2) \leq \alpha \leq 1/66$. Par conséquent, nous avons

$$b = \frac{2}{\alpha}\ln\left(\frac{2C}{\widehat{\alpha}\varepsilon}\right) \geq 132\ln\left(\frac{1169740}{3}\left(\frac{6544}{9}\right)^{18}\right) \geq 17354.$$

Nous pouvons donc dire que les constantes a et b obtenues à partir de [10] vérifient $a \geq 132$ et $b \geq 17354$, constantes qui sont d'au moins deux ordres de grandeur plus grandes que celles que nous avons obtenues (voir tableau 6.1).

6.5 Simulations

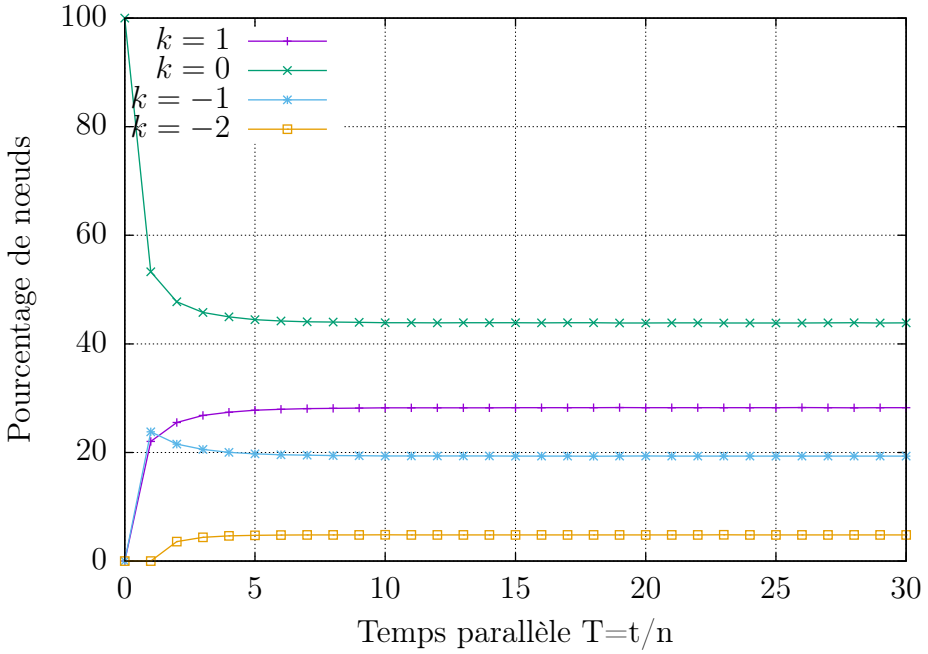


FIGURE 6.1 : Pourcentages moyens $Y_T(n,k) \times 100$ de nœuds à la valeur $T+k$, en fonction du temps parallèle T, pour $n = 1000$, et $k = -2, -1, 0, 1$.

Nous terminons ce chapitre en donnant un résumé des expérimentations que nous avons effectuées pour illustrer les performances de notre protocole. Rappelons que n est le nombre de nœuds dans le système et que $T = t/n$ est le nombre total d'interactions divisé par n, que l'on nomme temps parallèle. Nous avons conduit deux types d'expérimentations.

Le premier type illustre la proportion attendue de nœuds $Y_T(n,k)$ dont le compteur est égal à $T+k$ au temps nT, pour différentes valeurs de n et k. Plus précisément, $Y_T(n,k)$ est défini par

$$Y_T(n,k) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{C_{nT}^{(i)}=T+k\}}.$$

Dans la figure 6.1, nous montrons la valeur attendue de $Y_T(n,k)$, pour $n = 1000$ et $k = -2, -1, 0, 1$, en fonction du temps parallèle T. Ces résultats ont été obtenus

k \ n	10^3	10^4	10^5	10^6	10^7
-13	0.0	0.0	0.0	1.4E-9	1.42E-9
-12	0.0	2.0E-8	8.0E-9	9.0E-9	6.14E-9
-11	2.0E-7	4.0E-8	2.2E-8	2.8E-8	3.048E-8
-10	2.0E-7	8.0E-8	1.88E-7	1.436E-7	1.4814E-7
-9	4.0E-7	8.0E-7	7.7E-7	7.438E-7	7.2784E-7
-8	3.0E-6	3.6E-6	3.586E-6	3.48E-6	3.6029E-6
-7	1.42E-5	1.8E-5	1.8222E-5	1.7767E-5	1.7758E-5
-6	8.98E-5	8.602E-5	8.7176E-5	8.7372E-5	8.72753E-5
-5	4.372E-4	4.2706E-4	4.2957E-4	4.2901E-4	4.29349E-4
-4	0.0021144	0.0021023	0.0021071	0.0021092	0.0021086
-3	0.0102474	0.0102890	0.0102777	0.0102800	0.0102810
-2	0.0481626	0.0483366	0.0483382	0.0483465	0.0483437
-1	0.1930704	0.1932864	0.1933165	0.1933143	0.1933182
0	0.4389352	0.4380932	0.4380715	0.4380374	0.4380346
1	0.2824746	0.2827344	0.2826797	0.2827057	0.2827070
2	0.0243744	0.0245499	0.0245973	0.0245953	0.0245949
3	7.6E-5	7.224E-5	7.2248E-5	7.27752E-5	7.27974E-5
4	0.0	0.0	0.0	4.0E-10	3.6E-10

Tableau 6.2 : Espérance de $Y_{50}(n,k)$ en fonction du nombre de nœuds n et du décalage k

après avoir exécuté 10^4 expérimentations indépendantes. La figure 6.1 montre que la valeur attendue de $Y_T(n,k)$ semble converger lorsque T tend vers l'infini et que cette convergence est atteinte très rapidement. Notons que pour les autres valeurs de k, les proportions sont trop proches de 0 pour être représentées, comme indiqué dans le tableau 6.2 qui montre la proportion moyenne de nœuds $Y_T(n,k)$ dont le compteur est égal à $T+k$ au temps $T=50$, pour différentes valeurs de $n=10^3,10^4,10^5,10^6,10^7$ et $k=-13,\ldots,4$. Ces résultats ont été obtenus après avoir exécuté 5000 expérimentations indépendantes, pour chaque valeur de n. La valeur attendue de $Y_{50}(n,k)$ semble être presque indépendante de n pour les grandes valeurs de n.

Le second type d'expérimentations illustre le gap (avec les valeurs maximales, moyennes et minimales) pour différentes valeurs de la taille n du système. Soit $B=2\times10^9$ le nombre total d'interactions considérées. Le gap maximal est calculé comme suit : $\max_{100n\le t\le B}\mathrm{Gap}(t)$, le gap minimal est donné par $\min_{100n\le t\le B}\mathrm{Gap}(t)$, et le gap moyen est donné par

$$\frac{1}{B-100n}\sum_{t=100n}^{B-1}\mathrm{Gap}(t).$$

La figure 6.2 montre respectivement le gap minimal, moyen et maximal dans un système de taille n sur l'intervalle de temps $[100n,B]$. Rappelons que le temps est mesuré par le nombre d'interactions. Comme on pouvait s'y attendre, la progression logarithmique du gap est manifeste.

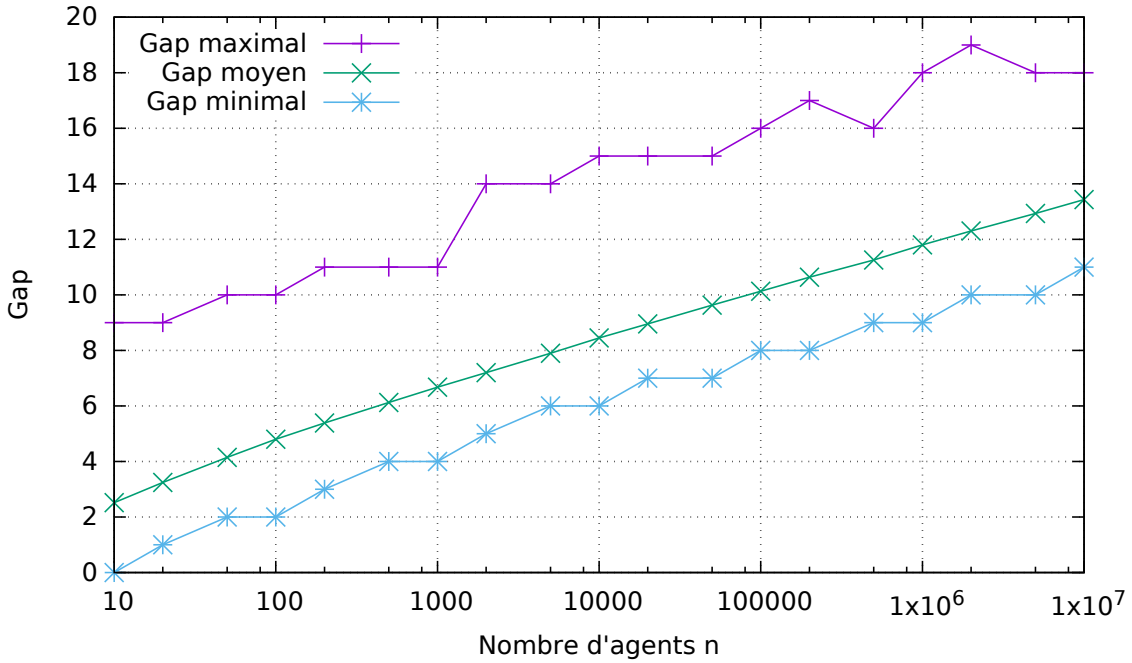


FIGURE 6.2 : Gap maximal, moyen et minimal en fonction de n.

6.6 Conclusion

Dans ce chapitre, nous avons fait progresser l'étude du paradigme "two-choice" en fournissant une analyse précise du problème du gap. Une application importante de cette étude serait l'amélioration des protocoles de population sans leader. En effet, nous avons montré que les agents peuvent disposer d'une horloge globale en garantissant que les valeurs de tous les compteurs d'agents sont concentrées selon la relation (6.1). Ils peuvent donc utiliser localement cette horloge globale pour déterminer les instants auxquels certaines actions spécifiques doivent être déclenchées, ou les instants à partir desquels tous les agents du système ont convergé vers un état donné. Dans le premier cas, cela permet aux agents de résoudre des problèmes plus complexes en déclenchant une série de protocoles de population, alors que dans le second cas, les agents peuvent déterminer l'instant à partir duquel tous les agents ont réussi à calculer une caractéristique donnée de la population. C'est d'ailleurs l'objet du chapitre suivant.

Chapitre 7

Détection de convergence

Ce chapitre propose un mécanisme qui permet à chaque agent de détecter localement que le système a convergé vers la configuration recherchée avec une probabilité élevée. Pour illustrer notre mécanisme, nous l'utilisons pour détecter l'instant auquel le protocole de proportion (déjà traité en section 5.2) converge. Spécifiquement, soit n_A (respectivement n_B) le nombre d'agents initialement démarrés dans l'état A (respectivement B) et $\gamma_A = n_A/n$, où n est le nombre total d'agents, notre protocole garantit, avec une précision donnée $\varepsilon > 0$ et une probabilité élevée $1 - \delta$, qu'après $O(n \ln(n/\delta))$ interactions, tout agent interrogé ayant le signal de détection positionné donnera la valeur correcte de la proportion γ_A des agents qui ont démarré à l'état A, avec un nombre d'états égal à $O(\ln(n)/\varepsilon)$. Pour autant que nous le sachions, nous n'avons pas connaissance de tels résultats dans d'autres études. Des résultats d'expérimentations illustrent notre analyse théorique.

7.1 Introduction

Dans ce chapitre, nous proposons un mécanisme permettant à chaque agent de détecter localement que le système a convergé vers la configuration recherchée avec une forte probabilité. En tant qu'application, nous proposons d'utiliser ce mécanisme pour détecter l'instant auquel le problème de proportion est résolu. Spécifiquement, soit n_A (respectivement n_B) le nombre d'agents initialement dans l'état A (respectivement B) et $\gamma_A = n_A/n$ (respectivement $\gamma_B = n_B/n$) la proportion d'agents dans l'état A (respectivement B) par rapport au nombre n d'agents. Notre protocole garantit, avec une précision donnée $\varepsilon > 0$ et n'importe quelle probabilité $1 - \delta$, que si un agent a détecté la convergence, alors chaque nœud peut calculer la valeur correcte de la proportion γ_A des agents qui ont démarré dans l'état A. De plus après $O(n \log(n/\delta))$ interactions, tout agent interrogé aura détecté la convergence, ceci avec un nombre d'états égal à $O(\log(n)/\varepsilon)$ (le nombre d'états de la proportion sans détection de convergence est de $O(1/\varepsilon)$).

Pour permettre à chaque nœud de détecter localement que le calcul de la proportion a convergé, nous combinons trois algorithmes, chacun étant exécuté sur chaque nœud du système. Le premier est dédié au calcul de la propriété recherchée, soit le calcul de la proportion γ_A. Le deuxième algorithme, exécuté en parallèle, maintient les données de l'horloge globale du système pour détecter l'instant auquel la

convergence est atteinte. En bref, lorsque l'horloge locale d'au moins deux nœuds a atteint un seuil donné $T_{\max}$, cela signifie que le nombre d'interactions globales dans le système est suffisamment important pour que tous les nœuds du système puissent calculer la proportion avec une grande précision. Ainsi, les deux nœuds peuvent commencer la propagation d'un signal pour informer les autres nœuds, qu'ils peuvent, à partir de leur état local, calculer une approximation de la proportion γ_A correcte à ε près. Cette dissémination est réalisée par le troisième algorithme qui a les même propriétés que la diffusion de rumeur (voir chapitre 4). Nous fournissons dans la section 7.4, un rappel de théorèmes ou une nouvelle formulation de théorèmes déjà existant. Comme cela sera également démontré dans la section 7.4, ce mécanisme de détection est universel en ce sens que tout protocole de population peut être complété par ce mécanisme pour détecter en toute sécurité la convergence avec une probabilité élevée. La seule exigence à satisfaire est qu'une borne supérieure du temps de convergence avec probabilité élevée pour ce protocole soit explicitement connue. Le nombre d'états est multiplié par cette borne, en temps parallèle.

Le reste du chapitre est organisé comme suit. La formalisation du problème, le modèle du système ainsi que les différentes notations adoptées dans le chapitre sont présentés à la section 7.2. L'orchestration des différents ingrédients de notre mécanisme de détection est présentée à la section 7.3. Une analyse théorique approfondie des performances de notre mécanisme de détection est présentée à la section 7.4, et les résultats des expérimentations sont présentés à la section 7.5. Enfin, la section 7.6 conclut le chapitre.

7.2 Modèle et notations

Dans ce travail, nous supposons que la collection de nœuds communiquent par des interactions par paires et de manière asynchrone. Initialement, chaque nœud commence avec un symbole initial A ou B. La fonction d'entrée ι initialise son état local en fonction de son symbole initial, puis à chaque interaction, son état est mis à jour en utilisant une fonction de transition notée f.

Les interactions entre les nœuds sont aléatoires : à chaque instant discret, deux agents quelconques sont choisis au hasard pour interagir. La notion de temps dans les protocoles de population désigne les étapes successives auxquelles les interactions se produisent, tandis que le temps parallèle désigne le nombre d'interactions moyen dont chaque nœud est l'initiateur, c'est à dire le nombre total d'interactions divisé par n, voir Aspnes et al. [15]. Rappelons que les nœuds ne gèrent ni n'utilisent d'identificateurs, cependant pour facilité la présentation, ils sont numérotés de 1 à n.

On note par le triplet $(C_t^{(i)}, T_t^{(i)}, S_t^{(i)})$ l'état du nœud i à l'instant t où $C_t^{(i)}$ est utilisé pour évaluer la valeur de la proportion du nœud i, $T_t^{(i)}$ représente la valeur de l'horloge globale du système et $S_t^{(i)}$ est une variable booléenne indiquant si la convergence des proportions a été atteinte globalement ou non.

Soient m et $T_{\max}$ deux paramètres du système. m est utilisé pour définir la configuration initiale des nœuds pour le calcul de la proportion. $T_{\max}$ détermine le nombre global d'interactions après lequel la convergence est atteinte pour tous les

nœuds. Les valeurs des deux paramètres sont analysées dans la section 7.4. À tout moment t, l'ensemble des états de $C_t^{(i)}$ est défini par $Q_C = [\![-m, m]\!]$ et nous avons donc $|Q_C| = 2m+1$; l'ensemble des états de $T_t^{(i)}$ est défini par $Q_T = [\![0, T_{\max} - 1]\!]$ et nous avons $|Q_T| = T_{\max}$; enfin l'ensemble des états de $S_t^{(i)}$ est défini par $Q_S = \{0, 1\}$, et nous avons $|Q_S| = 2$. Ainsi, l'ensemble des états d'un nœud est $Q = Q_C \times Q_T \times Q_S$ et donc la taille de cet ensemble est $|Q| = |Q_C \times Q_T \times Q_S| = |Q_C| \times |Q_T| \times |Q_S| = 2(2m+1)T_{\max}$.

La configuration du système à l'instant t est l'état de chaque nœud à l'instant t et est noté par le triplet (C_t, T_t, S_t) où $C_t = (C_t^{(1)}, \ldots, C_t^{(n)})$, $T_t = (T_t^{(1)}, \ldots, T_t^{(n)})$, et $S_t = (S_t^{(1)}, \ldots, S_t^{(n)})$.

Les interactions entre les nœuds sont orchestrées par un ordonnanceur aléatoire : à chaque instant discret $t \geq 0$, deux indices quelconques i et j sont choisis aléatoirement pour interagir avec probabilité $p_{i,j}(t)$. Les choix successifs du couple de nœuds en interaction sont supposés indépendants et uniformément répartis, ce qui signifie que nous avons

$$p_{i,j}(t) = \frac{1_{\{i \neq j\}}}{n(n-1)}.$$

7.3 Algorithme du protocole

Chaque nœud i maintient, comme état courant, un triplet $(C^{(i)}, T^{(i)}, S^{(i)})$, initialisé selon l'Algorithme 1.

```
1 Fonction ι(i, Symbole) :
2     si Symbole = A alors C^(i) := m ;
3     si Symbole = B alors C^(i) := −m ;
4 Fonction ι′(i, Symbole) :
5     ι (i, Symbole) ;
6     T^(i) := 0 ;
7     S^(i) := 0 ;
```

Algorithme 1 : Fonction d'entrée ι', initialisation du nœud i

Les couples de nœuds interagissent selon l'algorithme 2 et, pendant l'interaction, mettent à jour leur état en calculant la moyenne de leur valeur C et en incrémentant leur valeur d'horloge T, le tout en suivant respectivement les algorithmes 3 et 4. La fonction de transition f de l'algorithme de proportion est donnée par $f(x, y) = (\lfloor (x+y)/2 \rfloor, \lceil (x+y)/2 \rceil)$, qui correspond à la fonction `Moyenne` de l'algorithme 3. La fonction de transition du protocole de détection de convergence pour la proportion est donnée par la fonction `MajEtat` de l'algorithme 2. Dans les algorithmes, nous utilisons les indices i et j pour distinguer le nœud initiateur et le nœud récepteur. Il est important de garder à l'esprit que les nœuds n'ont aucune connaissance de ces indices, de même que les nœuds n'ont aucune connaissance directe du temps. C'est d'ailleurs pour cette dernière raison que dans les algorithmes, le temps, c'est-à-dire le nombre total d'interactions depuis l'origine représenté par l'indice t, n'est pas noté. Lorsque deux nœuds en interaction ont en même temps leur horloge égale à $T_{\max} - 1$,

cela signifie que le nombre d'interactions globales dans le système est suffisamment grand pour permettre à tous les nœuds du système de calculer localement la valeur de la proportion avec la précision ε, avec une probabilité supérieure à $1 - \delta$. Ils positionnent tous les deux leur valeur de signal à 1, ce qui est le point de départ de la diffusion (voir Algorithme 5). Si pendant une interaction, au moins un des deux nœuds a sa valeur de diffusion égale à 1, c'est-à-dire $S = 1$, alors les deux nœuds positionnent leur valeur S à 1, et de plus, les valeurs C et T ne sont plus mises à jour, et à toutes les interactions ultérieures, chacun de ces deux nœuds ne fait que "propager" le signal $S = 1$.

```
1 Fonction MajEtat(i, j) :
2 |  si Diffusion (i, j) = 0 alors
3 |  |  si Horloge (i, j) = 0 alors
4 |  |  |  Moyenne (i, j) ;
5 |  |  fin
6 |  fin
```

Algorithme 2 : Mise à jour de l'état des nœuds i et j pendant leur interaction

```
1 Fonction Moyenne(i,j) :
2 |  (C(i), C(j)) := ( ⌊(C(i)+C(j))/2⌋ , ⌈(C(i)+C(j))/2⌉ ) ;
```

Algorithme 3 : Traitement moyenne, mise à jour des valeurs de C des nœuds i et j qui interagissent

```
1 Fonction Horloge(i, j) :
2 |  si T(i) = T(j) = (T_max − 1) alors
3 |  |  S(i) := S(j) := 1 ;
4 |  |  retourner 1 ;
5 |  fin
6 |  si T(i) ≤ T(j) alors T(i) := T(i) + 1 ;
7 |  sinon T(j) := T(j) + 1 ;
8 |  retourner 0 ;
```

Algorithme 4 : Traitement horloge, mise à jour des valeurs T et éventuellement de S des nœuds i et j qui interagissent

Dans le cadre du protocole de proportion, quand on interroge le nœud i, il retourne son estimation ω_A de la proportion initiale de A en fonction de l'état courant $C^{(i)}$. Nous avons

$$\omega_A(C^{(i)}) = (m + C^{(i)})/(2m)$$

Dans le cadre du protocole de proportion avec détection de convergence, notons qu'en plus de la proportion, le nœud i retourne aussi la valeur du signal $S^{(i)}$ (voir

```
1 Fonction Diffusion(i, j) :
2 |   si S_t^(i) = S_t^(j) = 0 alors retourner 0 ;
3 |   S^(i) := S^(j) := 1 ;
4 |   retourner 1 ;
```

Algorithme 5 : Traitement diffusion, mise à jour des valeurs de S des nœuds i et j qui interagissent

Algorithme 6), ainsi nous avons

$$\omega'_A(C^{(i)}) = \left((m + C^{(i)})/(2m), S^{(i)}\right).$$

Comme il est démontré dans la section 7.4, si $S^{(i)} = 1$ alors la proportion calculée par le nœud i est une approximation de γ_A à ε près, avec une probabilité supérieure à $1 - \delta$. Notons que si $S^{(i)} = 0$ alors nous ne savons pas si la proportion calculée par le nœud i est éloignée ou pas de γ_A.

```
1 Fonction ω'_A(i) :
2 |   retourner ( (m+C^(i))/(2m), S^(i) ) ;
```

Algorithme 6 : Fonction de sortie ω'_A pour le noeud i du protocole avec détection de convergence

7.4 Analyse

Cette section est consacrée à l'analyse de notre solution. Nous avons divisé l'analyse en cinq parties, la première étant consacrée à l'analyse de la fonction de diffusion de rumeurs (voir la section 7.4.1), la seconde à l'analyse de la fonction moyenne (voir la section 7.4.2) et le troisième à la fonction d'horloge globale (voir la section 7.4.3). Il faut noter que, pour ces trois premières sections, il s'agit essentiellement d'un rappel ou d'une mise en forme de lemmes ou de théorèmes déjà démontrés dans les chapitres précédents. L'analyse présentée dans la quatrième section consiste à évaluer le comportement global de notre protocole en combinant les évaluations précédentes (voir la section 7.4.4). Nous terminons cette section en montrant sous quelle hypothèse notre mécanisme de détection de convergence peut être appliqué à tout protocole basé sur l'interaction par paires (voir la section 7.4.5).

7.4.1 Analyse du protocole de diffusion

La diffusion commence au premier instant t auquel deux nœuds en interaction, disons i et j, ont leur valeur de diffusion, $S_t^{(i)}$ et $S_t^{(j)}$, égale à 1. Cet instant se produit lorsque les deux nœuds ont leur horloge égal à $T_{\max} - 1$ (c'est-à-dire $T_t^{(i)} = T_t^{(j)} = T_{\max} - 1$). Afin d'analyser le temps de convergence nous utilisons le théorème 4.2.5 démontré au chapitre 4, nous allons le rappeler en modifiant légèrement la notation. Soit Y_t

le nombre de nœuds informés à l'instant t et $\mathcal{T}_n$[1] le premier instant auquel tous les nœuds connaissent la rumeur. Nous avons

$$\mathcal{T}_n = \inf\{t \geq 0 \mid Y_t = n\}.$$

Notons que, dans notre cas, nous avons $Y_0 = 2$. Ce théorème donne une valeur maximale du temps de propagation avec une probabilité inférieure ou égale à une probabilité fixe $\delta \in]0,1[$ lorsque le système commence initialement avec deux nœuds informés.

Théorème 4.2.5 Pour tout $\delta \in]0,1[$, nous avons

$$\mathbb{P}\left\{\mathcal{T}_n \geq n\left(\ln(n) - \ln(\delta)/2\right) \mid Y_0 = 2\right\} \leq \delta.$$

Comme nous l'avons noté en section 4.2.3.3 et montré à l'aide d'expérimentations en section 4.2.5, ce théorème donne une valeur extrêmement précise du temps de convergence.

7.4.2 Analyse du protocole de proportion

Le protocole de proportion fonctionne avec le vecteur C_t. Il a été analysé en détail dans la section 5.2.3, nous rappelons la définition de sa fonction de transition f, pour deux nœuds interagissant i et j

$$\left(C_{t+1}^{(i)}, C_{t+1}^{(j)}\right) = f\left(C_t^{(i)}, C_t^{(j)}\right) = \left(\left\lfloor \frac{C_t^{(i)} + C_t^{(j)}}{2} \right\rfloor, \left\lceil \frac{C_t^{(i)} + C_t^{(j)}}{2} \right\rceil\right) \tag{7.1}$$

$$\text{et} \quad C_{t+1}^{(r)} = C_t^{(r)} \text{ pour } r \neq i, j.$$

et sa fonction de sortie

$$\omega_A(C_t^{(i)}) = (m + C_t^{(i)})/(2m).$$

La constante m est entière avec $m \geq 2$ et elle dépend de la précision souhaitée. Le protocole de proportion est défini par $\mathcal{P} = (\Sigma, Q_C, \Xi, \iota, f, \omega_A)$. Et nous avons $\Sigma = \{A, B\}$, $Q_C = [\![-m, m]\!]$, $\Xi = \omega_A(Q_C)$, $\iota(A) = m$ et $\iota(B) = -m$. f et ω_A ont été définis précédemment.

Nous rappelons aussi le théorème 5.2.13 qui donne son temps de convergence.

Théorème 5.2.13 Pour tout $\delta \in]0,1[$ et pour $\varepsilon \in]0,1[$, en prenant $m = \lceil 3/(4\varepsilon) \rceil$, nous avons, pour tout $t \geq \tau$,

$$\mathbb{P}\{|\omega_A(C_t^{(i)}) - \gamma_A| < \varepsilon \text{ pour tout } i \in [\![1, n]\!]\} \geq 1 - \delta,$$

où

$$\tau = n\left(3.12 \ln n - 2 \ln \varepsilon - 6.59 \ln \delta + 1.88\right).$$

Nous rappelons aussi ce lemme.

Lemme 5.2.14 Soient $\varepsilon \in]0,1[$ et l'événement E_t défini par

$$E_t = \left\{\left|\omega_A\left(C_t^{(i)}\right) - \gamma_A\right| < \varepsilon \text{ pour tout } i \in [\![1, n]\!]\right\},$$

alors la suite (E_t) est croissante.

[1] Dans la mesure où le temps est discret, par cohérence avec le chapitre 4 nous aurions dû appeler cet instant T_n, mais dans ce présent chapitre la lettre T est déjà utilisée.

7.4.3 Analyse du protocole d'horloge

Le protocole d'horloge fonctionne avec le vecteur T_t. Il a été analysé en détail dans le chapitre 6 et nous rappelons sa fonction de transition pour deux nœuds interagissant i et j.

$$\left(T_{t+1}^{(i)}, T_{t+1}^{(j)}\right) = \begin{cases} \left(T_t^{(i)} + 1, T_t^{(j)}\right) & \text{si } T_t^{(i)} \leq T_t^{(j)} \\ \left(T_t^{(i)}, T_t^{(j)} + 1\right) & \text{si } T_t^{(i)} > T_t^{(j)}. \end{cases} \quad \text{et} \quad T_{t+1}^{(r)} = T_t^{(r)} \text{ pour } r \neq i, j.$$

L'écart maximal entre les valeurs d'horloge de deux nœuds distincts à l'instant t est aussi appelé le gap à instant t. Il se note $\mathrm{Gap}(t)$ et est défini par

$$\mathrm{Gap}(t) = \max_{1 \leq i \leq n} \left(T_t^{(i)}\right) - \min_{1 \leq i \leq n} \left(T_t^{(i)}\right).$$

Le théorème suivant donne une valeur maximale du gap avec une probabilité fixée. Notons que cette valeur est indépendante du temps global t.

Théorème 7.4.1 *Pour tout $\delta \in]0,1[$, nous avons*

$$\mathbb{P}\left\{Gap(t) \geq 10\ln(n) - 10\ln(\delta) + 74\right\} \leq \delta.$$

Preuve. A partir du corollaire 6.3.14 et du tableau 6.1 dans lequel nous prenons $a = 10$ et $b = 74$, nous obtenons, pour tout $\sigma > 0$,

$$\mathbb{P}\left\{\mathrm{Gap}(t) \geq 10(1+\sigma)\ln(n) + 74\right\} \leq 1/n^{\sigma}.$$

Soit $\delta \in]0,1[$. En prenant $\sigma = -\ln(\delta)/\ln(n)$, nous obtenons $\sigma \ln(n) = -\ln(\delta)$ et $1/n^{\sigma} = \delta$, c'est-à-dire

$$\mathbb{P}\left\{\mathrm{Gap}(t) \geq 10\ln(n) - 10\ln(\delta) + 74\right\} \leq \delta,$$

ce qu'il fallait démontrer. ■

Voici deux propriétés qui sont utilisées dans la section suivante.

Comme, à chaque instant, l'horloge d'un et un seul nœud est incrémentée de 1, nous avons

$$\sum_{i=1}^{n} T_t(i) = t.$$

Il s'ensuit qu'à chaque instant $t \geq 0$, nous avons

$$\min_{1 \leq i \leq n} \left(T_t^{(i)}\right) \leq \frac{t}{n} \leq \max_{1 \leq i \leq n} \left(T_t^{(i)}\right). \tag{7.2}$$

7.4.4 Analyse du protocole de proportion avec détection de convergence

Nous combinons maintenant toutes les analyses précédentes pour étudier le comportement de notre protocole de proportion avec détection de convergence. Le protocole de proportion avec détection de convergence est défini par $\mathcal{P}' = (\Sigma, Q', \Xi', \iota', f', \omega'_A)$. Et nous avons, Σ qui reste le même que pour le protocole $\mathcal{P}$, $Q' = Q_C \times Q_T \times Q_S$, $\Xi' = \Xi \times Q_S$, ι' est définie dans l'algorithme 1, f' est la fonction de transition et elle correspond à la fonction `MajEtat` de l'algorithme 2, enfin ω'_A est définie dans l'algorithme 6.

Pour tout $n \geq 2$ et pour tout $\delta \in]0, 1[$, nous introduisons les constantes suivantes :

- $\tau_1 = \ln(n) - 0.5\ln(\delta) + 0.55$.
- $\tau_2 = \tau = 3.12\ln n - 2\ln\varepsilon - 6.59\ln\delta + 9.12$.
- $\tau_3 = 10\ln(n) - 10\ln(\delta) + 84.99$.

La constante τ_1 est la constante utilisée dans le théorème 4.2.5 (voir section 7.4.1) avec $\delta/3$ au lieu de δ. C'est un majorant du temps parallèle nécessaire au protocole de diffusion pour converger avec une probabilité supérieure à $1 - \delta/3$.

La constante τ_2 est la constante utilisée dans le théorème 5.2.13 (voir section 7.4.2) avec $\delta/3$ au lieu de δ. C'est un majorant du temps parallèle nécessaire au protocole de proportion pour converger avec une probabilité supérieure à $1 - \delta/3$.

La constante τ_3 est la constante utilisée dans le théorème 7.4.1 avec $\delta/3$ au lieu de δ. C'est le gap maximal du protocole d'horloge, obtenu avec une probabilité supérieure à $1 - \delta/3$.

Avec ces constantes, nous initialisons $T_{\max} = \tau_2 + \tau_3$. Le théorème suivant est le principal résultat de ce chapitre. Il établit que, après l'instant $n(T_{\max} + \tau_1)$, tous les nœuds peuvent calculer une approximation de γ_A à ε près et que la diffusion du signal de convergence est terminée, avec une probabilité plus grande que $1 - \delta$. Il établit aussi que, si à n'importe quel instant t, un nœud a sa valeur de diffusion égale à 1, alors tous les nœuds peuvent faire une approximation de γ_A à ε près, avec une probabilité plus grande que $1 - 2\delta/3$.

Pour simplifier l'écriture, nous introduisons la variable aléatoire Y_t, définie, pour tout $t \geq 0$, par

$$Y_t = \sum_{i=1}^{n} S_t^{(i)}.$$

Nous introduisons également les événements contraires E_t et $\overline{E}_t$, définis par

$$E_t = \left\{ \left| \omega_A(C_t^{(i)}) - \gamma_A \right| < \varepsilon, \forall i \in [\![1, n]\!] \right\}$$
$$\overline{E}_t = \left\{ \exists i \in [\![1, n]\!], \left| \omega_A(C_t^{(i)}) - \gamma_A \right| \geq \varepsilon \right\}.$$

Théorème 7.4.2 *Pour tout $\delta \in]0,1[$ et $t \geq n(T_{\max} + \tau_1)$, nous avons*

$$\mathbb{P}\left\{E_t, Y_t = n\right\} \geq 1 - \delta.$$

De plus, pour tout $\delta \in]0,1[$, nous avons

$$\mathbb{P}\left\{\forall\, t \geq 0, \{Y_t = 0\} \cup E_t\right\} \geq 1 - 2\delta/3.$$

Preuve. Le protocole de proportion et le protocole d'horloge démarrent à l'instant 0 et s'exécutent séparément mais sur les mêmes interactions. Par conséquent, les deux processus ne sont pas indépendants, mais nous allons les analyser séparément.

Considérons d'abord le protocole d'horloge. Soit Γ le premier instant auquel deux nœuds en interaction ont leur valeur d'horloge égale à $T_{\max} - 1$. En appliquant le théorème 7.4.1 à l'instant Γ avec $\delta/3$ au lieu de δ, nous avons

$$\mathbb{P}\left\{\mathrm{Gap}(\Gamma) \geq 10\ln(n) - 10\ln(\delta/3) + 74\right\} \leq \delta/3,$$

c'est-à-dire, par définition de τ_3, $\mathbb{P}\left\{\mathrm{Gap}(\Gamma) \geq \tau_3\right\} \leq \delta/3$. Par définition du gap et comme à l'instant Γ nous avons $\max_{1\leq i\leq n}\left(T_\Gamma^{(i)}\right) = T_{\max} - 1$, nous obtenons

$$\mathbb{P}\left\{T_{\max} - 1 - \min_{1\leq i\leq n}\left(T_\Gamma^{(i)}\right) \geq \tau_3\right\} \leq \delta/3,$$

c'est-à-dire, par définition de $T_{\max}$,

$$\mathbb{P}\left\{\min_{1\leq i\leq n}\left(T_\Gamma^{(i)}\right) \leq \tau_2\right\} \leq \delta/3.$$

A partir de la relation (7.2) nous avons $\min_{1\leq i\leq n}\left(T_\Gamma^{(i)}\right) \leq \Gamma/n$, ce qui mène à

$$\mathbb{P}\left\{\Gamma \leq n\tau_2\right\} \leq \delta/3. \tag{7.3}$$

Considérons maintenant le protocole de proportion. En appliquant le théorème 5.2.13 rappelé en section 7.4.2 avec $\delta/3$ au lieu de δ, nous obtenons, par définition de τ_2, pour tout $t \geq n\tau_2$,

$$\mathbb{P}\left\{\overline{E}_t\right\} \leq \delta/3. \tag{7.4}$$

Puisque d'un point de vu logique $A \cup B = (A \cap \overline{B}) \cup B$, en utilisant les inégalités (7.3) et (7.4) d'une part, et la décroissance de $\overline{E}_t$ (voir lemme 5.2.14) d'autre part, nous avons, pour tout $t \geq 0$,

$$\begin{aligned}
\mathbb{P}\left\{\overline{E}_{\Gamma+t} \cup \left\{\Gamma \leq n\tau_2\right\}\right\} &= \mathbb{P}\left\{\left(\overline{E}_{\Gamma+t}, \Gamma > n\tau_2\right) \cup \left\{\Gamma \leq n\tau_2\right\}\right\} \\
&\leq \mathbb{P}\left\{\overline{E}_{\Gamma+t}, \Gamma > n\tau_2\right\} + \mathbb{P}\left\{\Gamma \leq n\tau_2\right\} \\
&\leq \mathbb{P}\left\{\overline{E}_{n\tau_2+t}, \Gamma > n\tau_2\right\} + \mathbb{P}\left\{\Gamma \leq n\tau_2\right\} \\
&\leq \mathbb{P}\left\{\overline{E}_{n\tau_2+t}\right\} + \mathbb{P}\left\{\Gamma \leq n\tau_2\right\} \\
&\leq 2\delta/3.
\end{aligned}$$

Il s'ensuit que, pour tout $t \geq 0$,

$$\mathbb{P}\left\{\overline{E}_{\Gamma+t}\right\} \leq \mathbb{P}\left\{\overline{E}_{\Gamma+t} \cup \{\Gamma \leq n\tau_2\}\right\} \leq 2\delta/3. \tag{7.5}$$

Γ est un temps d'arrêt pour $(Y_t)_t$. Le point de départ de la diffusion est $\Gamma+1$. Par définition de Γ, l'instant $\Gamma+1$ est le premier instant où exactement 2 agents ont leur valeur de diffusion égale à 1. Nous avons $Y_t = 0$ pour tout $t \leq \Gamma$ et $Y_{\Gamma+1} = 2$. Notons que le processus $(Y_t)_t$ est croissant et que $(Y_t)_{t>\Gamma}$ est une chaîne de Markov. En effet, le point de départ de la diffusion dépend du processus horloge par l'intermédiaire de Γ, tandis que le processus de diffusion n'est pas influencé par les autres protocoles.

En appliquant le théorème 4.2.5 avec $\delta/3$ au lieu de δ, par définition de τ_1, et puisque $(Y_t)_{t>\Gamma}$ est homogène, pour tout $t \geq 0$, nous avons

$$\begin{aligned}\mathbb{P}\left\{Y_{\Gamma+1+n\tau_1+t} \neq n\right\} &= \mathbb{P}\left\{Y_{\Gamma+1+n\tau_1+t} \neq n \mid Y_{\Gamma+1} = 2\right\} \\ &= \mathbb{P}\left\{Y_{n\tau_1+t} \neq n \mid Y_0 = 2\right\} \\ &\leq \delta/3. \end{aligned} \tag{7.6}$$

Rappelons que $\Gamma = \sum_{i=1}^{n} T_\Gamma^{(i)}$ et que $T_\Gamma^{(i)} \leq T_{\max} - 1$, pour tout $i \in [\![1, n]\!]$. Nous obtenons donc $\Gamma + 1 \leq nT_{\max}$. Par conséquent, pour tout $t \geq 0$, nous avons

$$\mathbb{P}\left\{Y_{nT_{\max}+n\tau_1+t} \neq n\right\} \leq \mathbb{P}\left\{Y_{\Gamma+1+n\tau_1+t} \neq n\right\},$$

et de même

$$\mathbb{P}\left\{\overline{E}_{nT_{\max}+n\tau_1+t}\right\} \leq \mathbb{P}\left\{\overline{E}_{\Gamma+1+n\tau_1+t}\right\},$$

ce qui donne en utilisant (7.5) et (7.6), pour $t \geq 0$,

$$\begin{aligned}\mathbb{P}\left\{\left\{Y_{nT_{\max}+n\tau_1+t} \neq n\right\} \cup \overline{E}_{nT_{\max}+n\tau_1+t}\right\} &\leq \mathbb{P}\left\{Y_{nT_{\max}+n\tau_1+t} \neq n\right\} + \mathbb{P}\left\{\overline{E}_{nT_{\max}+n\tau_1+t}\right\} \\ &\leq \mathbb{P}\left\{Y_{\Gamma+1+n\tau_1+t} \neq n\right\} + \mathbb{P}\left\{\overline{E}_{\Gamma+1+n\tau_1+t}\right\} \\ &\leq \delta/3 + 2\delta/3 = \delta,\end{aligned}$$

ce qui est équivalent, pour tout $t \geq n(T_{\max} + \tau_1)$, à

$$\mathbb{P}\left\{Y_t = n, E_t\right\} \geq 1 - \delta,$$

ce qui termine la première partie de la preuve.

Pour la seconde partie, notons que, pour tout $t \geq 0$,

$$Y_t \neq 0 \Longleftrightarrow t > \Gamma.$$

Par conséquent, en utilisant la décroissance de $(\overline{E}_t)$ (voir lemme 5.2.14), en appliquant la relation (7.5) et puisque Γ est borné (par $nT_{\max}$), nous avons

$$\begin{aligned}\mathbb{P}\left\{\exists\, t \geq 0, \text{ tel que } Y_t \neq 0, \overline{E}_t\right\} &= \mathbb{P}\left\{\exists\, t \geq 0, \text{ tel que } t > \Gamma, \overline{E}_t\right\} \\ &\leq \mathbb{P}\left\{\overline{E}_{\Gamma+1}\right\} \\ &\leq 2\delta/3.\end{aligned}$$

D'autre part, d'un point de vue logique, nous avons

$$\overline{\left\{\exists\, t \geq 0,\ \text{tel que } Y_t \neq 0, \overline{E}_t\right\}} = \left\{\forall\, t \geq 0, \{Y_t = 0\} \cup E_t\right\}$$

Par conséquent

$$\mathbb{P}\left\{\forall\, t \geq 0, \{Y_t = 0\} \cup E_t\right\} \geq 1 - 2\delta/3,$$

ce qui prouve la deuxième et dernière partie du théorème.

■

Ce dernier théorème montre que le temps de convergence est $O(\ln(n))$ et que le nombre d'états nécessaires est égal à $|Q_C \times Q_T \times Q_S| = 2\left(2\lceil 3/(4\varepsilon)\rceil + 1\right) T_{\max} = O\left(\ln(n)/\varepsilon\right)$.

7.4.5 Généralisation du mécanisme de détection de convergence

Maintenant, nous montrons que notre mécanisme de détection peut être appliqué à n'importe quel protocole $\mathcal{P} = (\Sigma, Q_C, \Xi, \iota, f, \omega)$ basé sur l'interaction par paires de sorte que tout nœud du système puisse détecter en toute sécurité l'instant auquel la convergence est atteinte par tous les nœuds du système. La seule exigence pour que ce mécanisme puisse s'appliquer est que le temps de convergence avec probabilité élevée de $\mathcal{P}$ doit être explicitement connu.

Spécifiquement, considérons la fonction de transition f du protocole $\mathcal{P}$ tel que la relation (7.1) soit remplacée par

$$\left(C_{t+1}^{(i)}, C_{t+1}^{(j)}\right) = f\left(C_t^{(i)}, C_t^{(j)}\right) \quad \text{et} \quad C_{t+1}^{(r)} = C_t^{(r)} \text{ pour } r \neq i, j.$$

La ligne 2 de l'algorithme 3 est donc remplacée par

$$\left(C^{(i)}, C^{(j)}\right) := f\left(C^{(i)}, C^{(j)}\right).$$

Les lignes 2 et 3 de la fonction ι de l'algorithme 1 sont remplacées par le corps de la fonction ι du protocole $\mathcal{P}$.

Comme nous l'avons fait en section 7.4.4 avec le protocole de proportion, à partir du protocole $\mathcal{P} = (\Sigma, Q_C, \Xi, \iota, f, \omega)$, nous définissons le nouveau protocole avec détection de convergence $\mathcal{P}' = (\Sigma, Q', \Xi', \iota', f', \omega')$. Σ reste le même que dans le protocole $\mathcal{P}$, $Q' = Q_C \times Q_T \times Q_S$, $\Xi' = \Xi \times Q_S$, ι' est définie dans l'algorithme 1 modifié précédemment, f' est la fonction de transition définie par la fonction `MajEtat` de l'algorithme 2, enfin ω' est définie pour $i \in [\![1, n]\!]$ et $t \geq 0$ par $\omega'\left(C_t^{(i)}, T_t^{(i)}, S_t^{(i)}\right) = \left(\omega\left(C_t^{(i)}\right), S_t^{(i)}\right)$.

La valeur initiale C_0 du vecteur C_t est donnée par la fonction d'entrée ι et l'état d'entrée de chaque noeud. L'ensemble des états Q_C de $C_t^{(i)}$ est supposé fini. Comme indicateur de convergence, nous considérons la fonction, spécifique au protocole $\mathcal{P}$, ν de $(Q_C)^n$ vers $\{0, 1\}$, cette fonction est croissante, ce qui signifie que si à un instant donné le protocole $\mathcal{P}$ converge, alors il converge aussi tous les instants suivants. Nous

supposons avoir une version spécifique au protocole $\mathcal{P}$ du théorème 5.2.13 établissant que pour tout $\delta \in]0,1[$ et pour tout $t \geq \tau_C(n,\delta)$, nous avons

$$\mathbb{P}\left\{\nu(C_t) = 1\right\} \geq 1 - \delta. \tag{7.7}$$

Sous les hypothèses précédentes, nous pouvons généraliser le théorème 7.4.2 de la manière suivante. Nous initialisons $T_{\max} = \tau_C(n,\delta/3) + \tau_3$.

Théorème 7.4.3 *Pour tout $\delta \in]0;1[$ et pour tout $t \geq n(T_{\max} + \tau_1)$ nous avons*

$$\mathbb{P}\left\{\nu(C_t) = 1, Y_t = n\right\} \geq 1 - \delta.$$

De plus, pour tout $\delta \in]0,1[$, nous avons

$$\mathbb{P}\left\{\forall\, t \geq 0, \{Y_t = 0\} \cup \{\nu(C_t) = 1\}\right\} \geq 1 - 2\delta/3.$$

Preuve. En définissant $E_t = \{\nu(C_t) = 1\}$, la preuve est exactement la même que celle du théorème 7.4.2, dans laquelle τ_2 est remplacé par $\tau_C(n,\delta/3)$. ■

Le nombre d'états dans ce cas est $|Q_C \times Q_T \times Q_S| = 2T_{\max}|Q_C|$.

7.5 Expérimentations

Dans cette section, nous présentons d'abord les résultats d'expérimentations pour les protocoles de diffusion, de proportion et d'horloge, puis nous présentons les résultats d'expérimentations pour le protocole complet.

7.5.1 Diffusion

Les résultats de la simulation de la diffusion ont déjà été présentés en section 4.2.5. Rappelons que les résultats de simulation sont remarquablement proches des résultats théoriques.

7.5.2 Proportion

Pour chaque valeur de ε, nous prenons $m = \lceil 3/(4\varepsilon) \rceil$. Ensuite nous choisissons $n_A = \lceil n/4m + n/2 \rceil$ et $n_B = n - n_A$. La raison de ce choix est d'obtenir, autant que cela est possible, une valeur de $\ell - \lfloor \ell \rfloor$ proche de 0.5, car, comme nous l'avons montré en section 5.2.6 avec la figure 5.4 d'une part et avec le théorème 5.2.10 d'autre part, la valeur de $\ell - \lfloor \ell \rfloor$ influence le temps de convergence, le pire des cas étant quand $\ell - \lfloor \ell \rfloor = 0.5$. Une simulation consiste à effectuer les interactions du protocole de proportion tel que décrit dans la section 7.4.2. La simulation s'arrête quand la différence entre les valeurs minimales et maximales des entrées du vecteur C_t devient inférieure ou égale à 2. N simulations indépendantes sont exécutées, les N valeurs du nombre d'interactions effectuées sont mémorisées puis ordonnées de cette manière : $\theta_1 \leq \theta_2 \leq \ldots \leq \theta_N$. L'estimation de l'instant τ tel que, pour $t \geq \tau$,

$$\mathbb{P}\left\{|\omega_A(C_t^{(i)}) - \gamma_A| < \varepsilon \text{ pour tout } i \in [\![1,n]\!]\right\} \geq 1 - \delta,$$

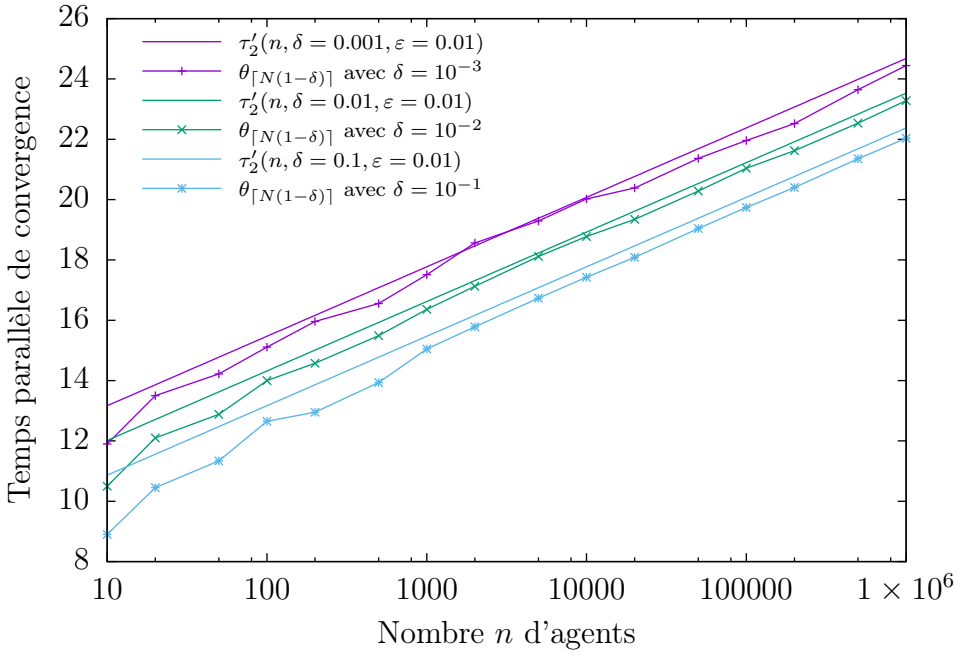


FIGURE 7.1 : Temps parallèle de convergence du protocole de proportion en fonction de n, avec $N = 10^4$ et $\varepsilon = 0.01$.

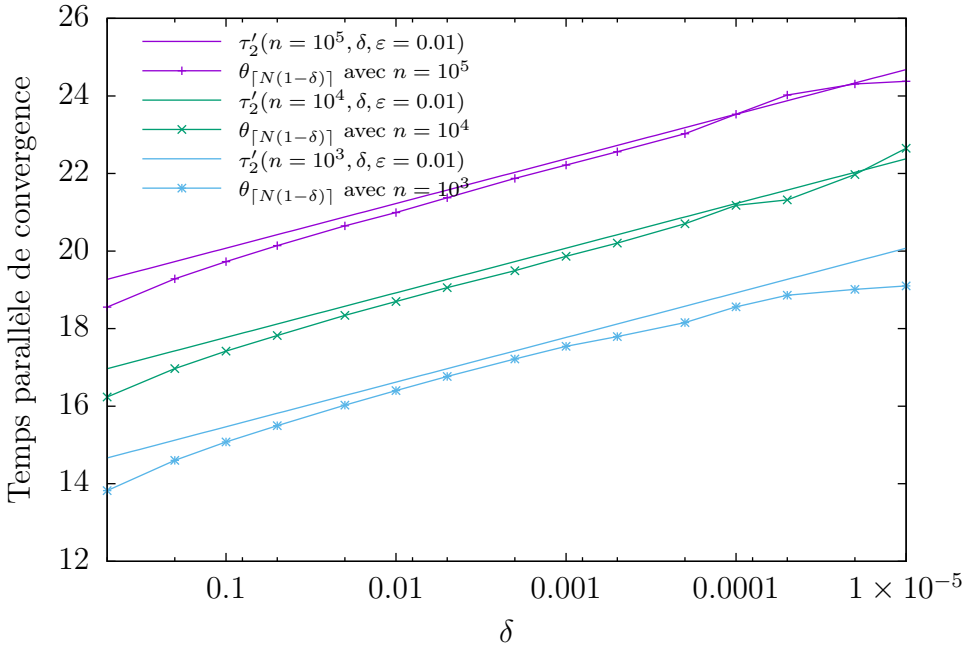


FIGURE 7.2 : Temps parallèle de convergence du protocole de proportion en fonction de δ, avec $N = 10^5$ et $\varepsilon = 0.01$.

est par conséquent donnée par la valeur $\theta_{\lceil N(1-\delta)\rceil}$.

Les figures 7.1, 7.2 et 7.3 représentent le temps parallèle de convergence $\theta_{\lceil N(1-\delta)\rceil}/n$ pour différentes valeurs de δ pour le premier, pour différentes valeurs de n pour le

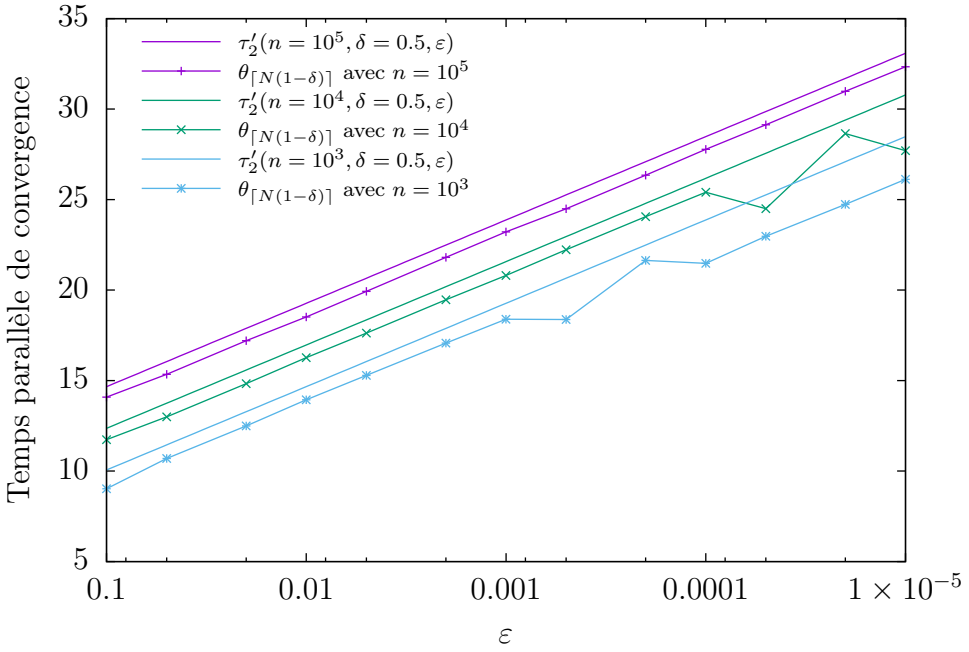


FIGURE 7.3 : Temps parallèle de convergence du protocole de proportion en fonction de ε, avec $N = 10^3$ et $\delta = 0.5$.

second et pour différentes valeurs de ε pour le dernier. Notons que, dans les deux premières figures, nous avons $\varepsilon = 0.01$, c'est-à-dire $m = \lceil 3/(4\varepsilon)\rceil = 75$. Dans chaque figure, les valeurs de $\theta_{\lceil N(1-\delta)\rceil}/n$ sont comparées à la valeur intuitive $\tau_2'(n, \delta, \varepsilon)$, proche de l'expression de τ_2 dont les coefficients sont déduits des résultats de simulation et sont donnés par

$$\tau_2'(n, \delta, \varepsilon) = \ln(n) - 0.5\ln(\delta) - 2\ln(\varepsilon) - 1.80. \tag{7.8}$$

Dans la figure 7.3, nous constatons un écart pour $n = 10^4$ et $n = 10^3$ quand $\varepsilon < 1/n$. Cet écart est dû au fait que la formule $n_A = \lceil n/4m + n/2\rceil$ ne permet plus d'obtenir $\ell - \lfloor \ell \rfloor = 0.5$.

7.5.3 Horloge

Pour le protocole d'horloge, une simulation consiste à effectuer les interactions du protocole d'horloge telles que décrites dans la section 7.4.3. Nous commençons l'évaluation du gap après les $50n$ premières interactions. A partir de ce moment, nous mémorisons le gap toutes les 100 interactions. x simulations sont exécutées indépendamment et pour chaque simulation le gap est mémorisé y fois. Cela veut dire que la durée d'une simulation est égale à $100y + 50n$. Le nombre N des valeurs de gap obtenues est donc $N = xy$. Ces N valeurs sont mémorisées et ordonnées de cette manière : $\text{Gap}_1 \leq \text{Gap}_2 \leq \ldots \leq \text{Gap}_N$. L'estimation de la valeur τ telle que

$$\mathbb{P}\left\{\text{Gap}(t) \geq \tau\right\} \leq \delta,$$

est donc donnée par la valeur $\text{Gap}_{\lceil N(1-\delta)\rceil}$.

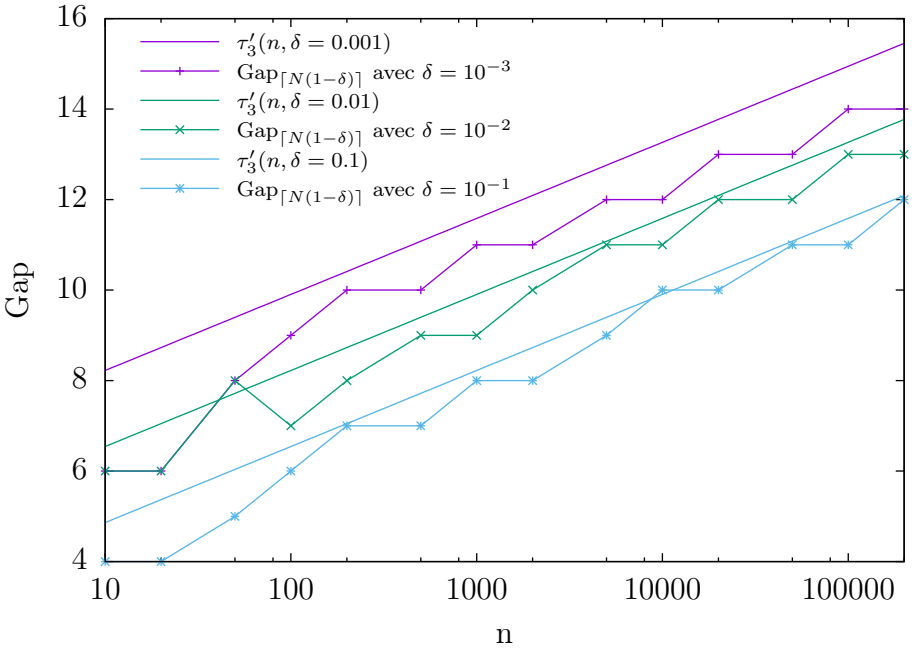


FIGURE 7.4 : Gap de l'horloge en fonction de n, avec $N = 10^6$.

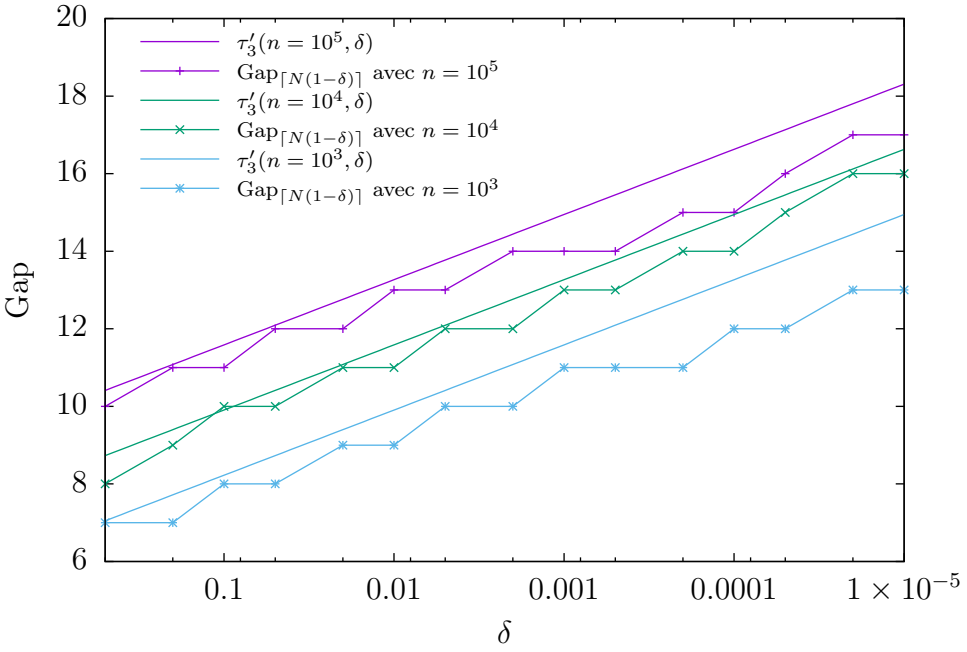


FIGURE 7.5 : Gap de l'horloge en fonction de δ, avec $N = 10^7$.

Les figures 7.4 et 7.5 décrivent la valeur $\mathrm{Gap}_{\lceil N(1-\delta)\rceil}$ du gap pour différentes valeurs de δ pour la première et pour différentes valeurs de n pour la seconde. Pour la figure 7.4 nous avons pris $x = 10$ et $y = 10000$ et pour la figure 7.5 nous avons pris $x = 100$ et $y = 10000$. Dans chaque figure, les valeurs de $\mathrm{Gap}_{\lceil N(1-\delta)\rceil}$ sont comparées à la valeur intuitive $\tau'_3(n, \delta)$ proche de l'expression de τ_3 dont les coefficients déduits

des résultats de simulation sont donnés par

$$\tau_3'(n,\delta) = 0.73\ln(n) - 0.73\ln(\delta) + 1.5. \tag{7.9}$$

7.5.4 Protocole optimisé déduit des expérimentations

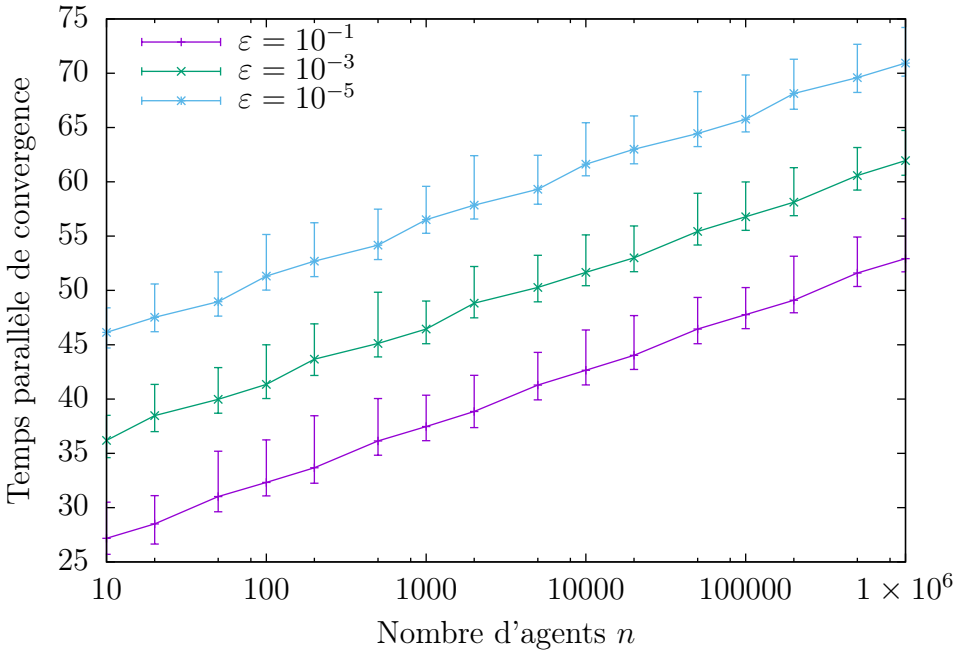


FIGURE 7.6 : Temps parallèle de convergence du protocole de proportion avec détection de convergence en fonction de n, avec $N = 10^3$.

Nous déduisons des relations (7.8) et (7.9) une valeur intuitive de $T'_{\max}$ proche de l'expression de $T_{\max}$ pour le protocole de proportion avec détection de convergence. Il est donné par

$$T'_{\max} = \tau_2'(n, \delta/3, \varepsilon) + \tau_3'(n, \delta/3) = 1.73\ln(n) - 1.23\ln(\delta) - 2\ln(\varepsilon) + 1.05.$$

Pour différentes valeurs de n et ε, $N = 1000$ simulations indépendantes on été exécutées en prenant $\delta = 10^{-6}$, en utilisant la valeur $T'_{\max}$ au lieu de $T_{\max}$. Les temps de convergence ont été mémorisés $\theta_1, \ldots, \theta_N$, définis, pour $i = 1, \ldots, N$, par

$$\theta_i = \inf\left\{t \geq 0 \mid S_t^{(j)} = 1, \forall j \in [\![1, n]\!]\right\}.$$

La figure 7.6 donne les résultats de simulation pour différente valeurs de ε. La valeur moyenne $(\theta_1 + \cdots + \theta_N)/N$, la valeur minimale $\min_{i=1,\ldots,N} \theta_i$ et la valeur maximale $\max_{i=1,\ldots,N} \theta_i$ y sont montrées. La leçon la plus importante que l'on peut tirer de toutes ces simulations est que le temps de convergence du protocole de proportion avec le mécanisme de détection de convergence est du même ordre de grandeur que le temps de convergence du protocole de proportion sans aucun mécanisme de détection. C'est un excellent résultat.

7.6 Conclusion

Dans ce chapitre, nous avons présenté comment améliorer, dans le modèle des protocoles de population, un protocole de proportion avec un mécanisme de détection de convergence pour permettre à chaque nœud du système de détecter localement l'instant auquel la convergence est atteinte avec un probabilité donnée. Les ingrédients de notre solution ont déjà été sujet à une analyse théorique en profondeur dans les chapitres précédents. Une analyse des données de simulations de ces protocoles a mené à un paramétrage très fin de notre protocole. Les résultats de la simulation montrent l'impact vraiment très faible de notre mécanisme de détection sur le temps de convergence du protocole de proportion. Nous avons également montré l'applicabilité de notre mécanisme de détection de convergence à de nombreux autres protocoles basés sur l'interaction par paires.

Chapitre 8

Conclusion et perspectives

8.1 Autres problèmes

Cette section présente des problèmes solubles avec un protocole de population, non traités dans cette thèse. Il semble difficile d'en faire l'impasse.

Dans la mesure où la synchronisation de protocoles demande une formulation explicite du temps de convergence avec probabilité élevée, l'explicitation de cette borne pour tous ces protocoles fait partie des perspectives.

Nous ferons la même remarque que dans le chapitre 3, les protocoles ne sont vus que de l'extérieur. De chaque protocole $(\Sigma, Q, \Xi, \iota, f, \omega)$, nous ne décrivons précisément que l'ensemble d'entrée Σ et l'ensemble de sortie Ξ. Nous évoquerons la fonction de sortie sans la définir, cette fonction de sortie est appliquée à l'état d'un nœud. Nous rappelons que les états de tous les nœuds du système, c'est-à-dire la configuration, sont définis par le vecteur $C_t = (C_t^{(1)}, C_t^{(2)}, \ldots, C_t^{(n)}) \in Q^n$.

8.1.1 L'élection de leader

L'élection de leader est un protocole dont le but est de désigner un et un seul "leader", les autres agents étant "follower", ceci en partant d'agents indifférenciés, c'est-à-dire que l'ensemble des symboles d'entrée est un singleton.

Le terme "leader" signifie en français "meneur". Ce terme est un peu trompeur dans la mesure où, si l'on reste dans le cadre des protocoles de population, l'agent désigné comme "leader" ne mène rien, il a juste un statut différent. Avoir un agent différent par nature peut permettre certaines applications, nous y reviendrons en section 8.1.4.

La fonction de sortie ω rend 0 si l'agent concerné est "follower" et rend 1 si l'agent concerné est "leader", donc l'ensemble de sortie est $\Xi = \{0, 1\}$.

Une fois désigné, le leader ne doit pas changer. Pour respecter ceci, le plus simple est qu'à l'origine, tous les agents soient "leader" et qu'un agent devenant "follower" ne puisse plus redevenir "leader". Autrement dit, pour un agent donné, la fonction de sortie ω doit être décroissante en fonction du temps.

Définition 8.1.1 *Soit $\delta \in]0, 1[$. Un protocole d'élection de leader converge à l'ins-*

tant $\tau \geq 0$ *avec la probabilité* $1-\delta$*, si, pour tout* $t \geq \tau$*,*

$$\mathbb{P}\left\{\sum_{i=1}^{n} \omega\left(C_t^{(i)}\right) = 1\right\} \geq 1-\delta$$

Il existe un protocole trivial d'élection de leader dont le principe est le suivant : à l'origine, tous les agents sont dans l'état 1, la fonction de transition f est telle que $f(1,1) = (1,0)$, et les autres interactions sont invariantes par f. Ce protocole possède un temps parallèle de convergence de $O(n)$ et bien sûr un nombre d'états constant égal à 2. En 2015, David Doty et David Soloveichik [32] ont, de manière surprenante, montré qu'avec un nombre d'états fini et ne dépendant pas de n, le temps parallèle de convergence avec probabilité élevée est $\Omega(n)$, c'est-à-dire qu'il n'est pas possible de faire mieux que $O(n)$ en terme de temps parallèle de convergence avec probabilité élevée.

Par conséquent, tous les protocoles proposés dans la suite de cette section ont un nombre d'états dépendant de n.

Ce problème a récemment connu des progrès impressionnants que nous pouvons résumer avec le tableau 8.1. Ce tableau est inspiré de [39] et enrichi avec la colonne "Pré-requis". Pour exprimer une borne avec probabilité élevée nous utilisons l'abréviation anglaise w.h.p. (with high probability).

Année	Article	Pré-requis	États	Temps parallèle	Type
2015	[7]	néant	$O(\log^3 n)$	$O(\log^3 n)$	espérance
				$O(\log^4 n)$	w.h.p.
2017	[5]	néant	$O(\log^2 n)$	$O(\log^{5.3} n \log\log n)$	espérance
				$O(\log^{6.3})$	w.h.p.
2017	[22]	néant	$O(\log^2 n)$	$O(\log^2 n)$	w.h.p.
2018	[6]	néant	$O(\log n)$	$O(\log^2 n)$	espérance
2018	[20]	néant	$O(\log n)$	$O(\log^2 n)$	w.h.p.
2018	[38]	junte	$O(\log\log n)$	$O(\log^2 n)$	w.h.p.
2018	[39]	junte	$O(\log\log n)$	$O(\log n \log\log n)$	espérance

Tableau 8.1 : Progrès récents sur l'élection de leader avec les protocoles de population

Il faut noter que dans [38, 39], les protocoles, bien que très performants, nécessitent la présence d'une junte.

8.1.2 L'élection de junte

8.1.2.1 Définition

Le problème d'élection de junte est proche de celui de l'élection de leader, au lieu d'avoir un "leader", nous en avons plusieurs. Ce nombre doit être borné par une fonction c explicitement fournie telle que $c(n) = o(n)$. Le nombre de leaders, à la convergence du protocole, doit être inférieur à $c(n)$. Il s'agit d'une borne supérieure, la borne inférieure étant égale à 1 (il faut qu'il y ait au moins un "leader"). L'élection de junte est moins contraignante que l'élection de leader. Certains protocoles

nécessitant un leader peuvent fonctionner également avec une junte, comme nous le verrons en section 8.1.4 avec l'horloge avec leader ou junte. D'autre part, comme nous l'avons vu, la présence d'une junte peut être un pré-requis pour l'élection de leader [38, 39].

Comme pour l'élection de leader, la fonction de sortie ω rend 1 quand elle est appliquée à l'état d'un agent "leader" et 0 quand il s'agit d'un "follower". Cette fonction de sortie appliquée à un agent doit être décroissante.

Définition 8.1.2 *Soit $\delta \in]0,1[$. Un protocole d'élection de junte converge à l'instant $\tau \geq 0$ avec la probabilité $1-\delta$, lorsque, pour tout $t \geq \tau$,*

$$\mathbb{P}\left\{1 \leq \sum_{i=1}^{n} \omega\left(C_t^{(i)}\right) \leq c(n)\right\} \geq 1-\delta.$$

8.1.2.2 État de l'art

En mars 2018, Gasieniec et Stachowiak [38] ont proposé un protocole pour l'élection d'une junte. La fonction c attachée à ce protocole est $c(n) = (n \log n)^{1/2}$. Le nombre d'états nécessaires est $O(\log \log n)$ et le temps parallèle de convergence avec probabilité élevée est de $O(\log n)$.

8.1.3 Le consensus

8.1.3.1 Définition

Le consensus consiste à ce que chaque nœud réponde de manière identique à une question. La réponse n'est pas nécessairement exacte, elle peut n'être qu'approchée. Ce qui prime est que tous les nœuds répondent de la même manière. Tout problème de majorité est un problème de consensus, l'exemple typique est ce protocole à trois états [68] dont la valeur exacte n'est garantie que lorsque $|n_A - n_B| = \Omega\left(\sqrt{n \log n}\right)$. Ce protocole donne une réponse, pas forcément la bonne, mais cette unanimité est plus importante que l'exactitude.

8.1.3.2 État de l'art

En 2017, Cordasco et Gargano [28] ont proposé un protocole de consensus donnant la proportion. Pour ce faire, ils se sont appuyés sur notre article [55]. Ils l'ont utilisé pour que tous les nœuds puissent rendre une même valeur approchée de la proportion.

8.1.4 L'horloge avec leader ou avec junte

Cette horloge a été décrite pour la première fois dans [12]. Son principe simplifié est que le leader diffuse la valeur de son compteur, quand cette diffusion finit par le toucher à nouveau, alors il incrémente son compteur et relance la diffusion de la nouvelle valeur. En pratique, quand deux followers se rencontrent, ils prennent la valeur la plus grande des deux et quand le leader rencontre un follower ayant la même valeur que la sienne, les deux valeurs s'incrémentent, les autres interactions n'ont pas

d'effet. L'unité de temps de cette horloge correspond au temps de diffusion, c'est à dire $\Theta(\log n)$ en temps parallèle. Dans l'article déjà cité, il est suggéré que l'utilisation d'une junte permet le même résultat. Ce qui a été fait et expliqué dans [38]. La même horloge avec junte est également utilisée dans [39].

8.1.5 La proportion en tant que résultat

Ce problème consiste à définir un protocole ayant un nombre fini d'états indépendant de n. Ce protocole converge vers une configuration où la proportion de nœuds dans un certain état reste stable [62, 23, 4]. Cette proportion tend vers la valeur calculée par le protocole lorsque la taille du système n tend vers l'infini. Bournez et al. [62] ont prouvé l'ensemble des nombres calculables de cette manière était exactement l'ensemble des nombres algébriques de $[0, 1]$. À notre connaissance le temps de convergence de ce type de protocole n'a pas été abordé.

8.2 Conclusion

Dans cette thèse, après une présentation des principaux protocoles de population existants, nous avons étudié en profondeur quatre protocoles : le protocole de diffusion, de moyenne avec des entiers, de moyenne avec des réels et le protocole d'horloge. Pour chacun de ces protocoles, nous avons mis en évidence des résultats inédits qui ont un grand intérêt pratique.

Pour le protocole de diffusion, nous avons démontré que la queue de distribution est bornée par une expression simple et précise. Pour le protocole de moyenne avec des entiers, nous avons démontré la convergence en $O(\log n)$ vers un état où l'écart maximal entre deux valeurs est 2. Pour le protocole de moyenne avec des réels, par l'utilisation de la norme 4, nous avons réduit considérablement les constantes des bornes de convergence. Pour le protocole d'horloge basé sur le "two-choice", nous avons explicité les constantes des bornes.

Enfin, en utilisant trois de ces protocoles (diffusion, moyenne avec entiers et horloge), nous avons construit un protocole avec détection de convergence, ce qui, à notre connaissance, n'a jamais été fait.

8.3 Perspectives

Dans la conclusion de chaque chapitre, nous avons évoqué des pistes concernant chaque protocole. Nous pouvons en donner d'autres :

- Dans le chapitre 4, nous avons généralisé le problème de la diffusion au temps continu. Il serait intéressant d'effectuer cette généralisation pour les autres protocoles. Une manière élégante de procéder, pour autant que ce soit possible, serait de montrer sous quelles hypothèses un résultat en temps discret est généralisable au temps continu.

- Dans le la section 3.1.2, nous avons évoqué des variantes du protocole de la diffusion de rumeur, variantes avec "étouffeurs" [29, 50]. Il pourrait être

intéressant de les étudier dans le cadre des protocoles de population, au moyen d'une chaîne de Markov.

- En utilisant un protocole similaire à la moyenne avec les entiers ou les réels et en ajoutant la valeur 1 à chaque interaction, nous pouvons construire une horloge qui, expérimentalement, montre des propriétés équivalentes à celles de l'horloge construite au chapitre 6. Tout reste à faire pour l'analyse de cette nouvelle horloge.

- Dans le chapitre 7, nous avons construit un protocole avec détection de convergence. Il faudrait aller plus loin et montrer comment concaténer des protocoles de manière générique. Cela permettrait, dans le cas d'un protocole d'élection de leader qui a besoin d'une junte [39], de ne réaliser qu'un seul protocole qui enchaînerait l'élection de junte puis l'élection de leader. Ce nouveau protocole aurait $O(\log n \log\log n)$ états et aurait un temps parallèle espéré de $O(\log n \log\log n)$ et ceci sans pré-requis.

- Notre mécanisme de détection de convergence nécessite une formule explicite sur le temps de convergence avec probabilité élevée. De nombreux protocoles parmi les plus performants sont fournis sans cette formule explicite prouvée. Cette explicitation est indispensable à la concaténation des protocoles évoquée plus haut.

- En couplant le protocole de proportion de la section 5.2 et un protocole dont la proportion est le résultat défini en section 8.1.5, nous pourrions développer une méthode pour calculer un nombre algébrique à la précision souhaitée.

- Le théorème 4.2.4 donnant un équivalent de la fonction de distribution du temps de convergence de la diffusion de rumeur est d'une très grande précision. Il a été trouvé pour le plus simple des protocoles, mais c'est un modèle à suivre pour d'autres types de protocoles tel que le protocole d'horloge two-choice ou le protocole basé sur la moyenne. C'est à partir de la connaissance précise des différents protocoles qu'il sera possible de les assembler de manière efficace et prouvée.

Bibliographie

[1] H. Acan, A. Collevecchio, A. Mehrabian, and W. Nick. On the push&pull protocol for rumour spreading. In *Proceedings of the ACM Symposium on Principles of Distributed Systems (PODC)*, 2015.

[2] M. Adler, P. Berenbrink, and K. Schröder. Analyzing an infinite parallel job allocation process. In *Proceedings of the European Symposium on Algorithms (ESA)*, 1998.

[3] M. Adler, S. Chakrabarti, M. Mitzenmacher, and L. Rasmussen. Parallel randomized load balancing. *Random Structures & Algorithms*, 13(2) :159–188, 1998.

[4] M. Albenque and L. Gerin. On the algebraic numbers computable by some generalized Ehrenfest urns. *Discrete Mathematics and Theoretical Computer Science*, 14(2) :271–284, 2012.

[5] D. Alistarh, J. Aspnes, D. Eisenstat, R. Gelashvili, and R. Rivest. Time-space trade-offs in population protocols. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2560–2579, 2017.

[6] D. Alistarh, J. Aspnes, and R. Gelashvili. Space-optimal majority in population protocols. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA, 2018.

[7] D. Alistarh and R. Gelashvili. Polylogarithmic-time leader election in population protocols. In *Proceedings, Part II, of the 42Nd International Colloquium on Automata, Languages, and Programming*, ICALP, 2015.

[8] D. Alistarh, R. Gelashvili, and M. Vojnovic. Fast and exact majority in population protocols. In *Proceedings of the 34th annual ACM symposium on Principles of Distributed Computing (PODC)*, pages 47–56, 2015.

[9] D. Alistarh, S. Gilbert, R. Guerraoui, and M. Zadimoghaddam. How efficient can gossip be ? (on the cost of resilient information exchange). In *Automata, Languages and Programming*, 2010.

[10] D. Alistarh, J. Kopinsky, J. Li, and G. Nadiradze. The power of choice in priority scheduling. In *Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC)*, 2017.

[11] D. Angluin, J. Aspnes, Z. Diamadi, M. Fischer, and R. Peralta. Computation in networks of passively mobile finite-state sensors. *Distributed Computing*, pages 235–253, Mar. 2006.

[12] D. Angluin, J. Aspnes, and D. Eisenstat. Fast computation by population protocols with a leader. *Distributed Computing*, 21(2) :183–199, 2008.

[13] D. Angluin, J. Aspnes, and D. Eisenstat. A simple population protocol for fast robust approximate majority. *Distributed Computing*, 20(4) :279–304, 2008.

[14] D. Angluin, J. Aspnes, D. Eisenstat, and E. Ruppert. The computational power of population protocols. *Distributed Computing*, 20(4) :279–304, 2007.

[15] J. Aspnes and E. Ruppert. An introduction to population protocols. *Bulletin of the European Association for Theoretical Computer Science, Distributed Computing Column*, 93 :98–117, 2007.

[16] Y. Azar, A. Z. Broder, A. R. Karlin, and E. Upfal. Balanced allocations (extended abstract). In *Proceedings of the ACM Symposium on Theory of Computing (STOC)*, 1994.

[17] P. Berenbrink, A. Czumaj, T. Friedetzky, and N. D. Vvedenskaya. Infinite parallel job allocation (extended abstract). In *Proceedings of the ACM Symposium on Parallel Algorithms and Architectures (SPAA)*, pages 99–108, 2000.

[18] P. Berenbrink, A. Czumaj, A. Steger, and B. Vöcking. Balanced allocations : The heavily loaded case. *SIAM Journal on Computing*, 35(6) :1350–1385, 2006.

[19] P. Berenbrink, R. Elsässer, T. Friedetzky, D. Kaaser, P. Kling, and T. Radzik. A population protocol for exact majority with $O(\log^{5/3} n)$ stabilization time and asymptotically optimal number of states. *CoRR*, 2018.

[20] P. Berenbrink, D. Kaaser, P. Kling, and L. Otterbach. Simple and Efficient Leader Election. In *1st Symposium on Simplicity in Algorithms (SOSA 2018)*, OpenAccess Series in Informatics (OASIcs), pages 9 :1–9 :11, 2018.

[21] N. Berger, C. Borgs, J. T. Chayes, and A. Saberi. On the spread of viruses on the internet. In *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2005.

[22] A. Bilke, C. Cooper, R. Elsässer, and T. Radzik. Brief announcement : Population protocols for leader election and exact majority with $O(\log^2 n)$ states and $O(\log^2 n)$ convergence time. In *Proceedings of the ACM Symposium on Principles of Distributed Computing*, PODC, 2017.

[23] O. Bournez, P. Chassaing, J. Cohen, L. Gerin, and X. Koegler. On the convergence of population protocols when population goes to infinity. *Applied Mathematics and Computation*, 215(4) :1340 – 1350, 2009. Physics and Computation.

[24] K. Censor-Hillel, B. Haeupler, J. Kelner, and P. Maymounkov. Global computation in a poorly connected world : Fast rumor spreading with no dependence on conductance. In *Proceedings of the Annual ACM Symposium on Theory of Computing (STOC)*, 2012.

[25] F. Chierichetti, S. Lattanzi, and A. Panconesi. Rumor spreading in social networks. *Theoretical Computer Science*, 412(24) :2602–2610, 2011.

[26] F. Comets, F. Delarue, and R. Schott. Information transmission under random emission constraints. *Environmental Modelling & Software*, 23(6) :973–1009, 2014.

[27] F. Comets, C. Gallesco, S. Popov, and M. Vachkovskaia. Constrained information transmission on Erdös-Rényi graphs. *Markov Processes and Related Fields*, 22 :111–138, 2016.

[28] G. Cordasco and L. Gargano. Space-optimal proportion consensus with population protocols. In *Stabilization, Safety, and Security of Distributed Systems*, 2017.

[29] D. Daley and D. G. Kendall. Stochastic rumours. *IMA Journal of Applied Mathematics*, 1(1) :42–55, 1965.

[30] S. Daum, F. Kuhn, and Y. Maus. Rumor spreading with bounded indegree. In *Proceedings of the International Colloquium on Structural Information and Communication Complexity (SIROCCO)*, 2016.

[31] A. Demers, M. Gealy, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, and D. Terry. Epidemic algorithms for replicated datbase maintenance. In *Proceedings of the ACM Syposium on Principles of Distributed Systems (PODC)*, 1987.

[32] D. Doty and D. Soloveichik. Stable leader election in population protocols requires linear time. *CoRR*, abs/1502.04246, 2015.

[33] M. Draief and M. Vojnovic. Convergence speed of binary interval consensus. *SIAM Journal on Control and Optimization*, 50(3) :1087–11097, 2012.

[34] U. Feige, D. Peleg, P. Raghavan, and E. Upfal. Randomized broadcast in networks. *Random Structures and Algorithms*, 1(4) :447–460, 1990.

[35] N. Fountoulakis and K. Panagiotou. Rumor spreading on random regular graphs and expanders. *Random Structures and Algorithms*, 43(2) :201–220, 2013.

[36] A. Frieze and G. Grimmet. The shortest-path problem for graphs with random arc-lengths. *Discrete Applied Mathematics*, 10(1) :57–77, 85.

[37] A. J. Ganesh. Rumour spreading on graphs. Technical report, 2015. `https://people.maths.bris.ac.uk/~maajg/teaching/complexnets/rumours.pdf`

[38] L. Gasieniec and G. Stachowiak. Fast space optimal leader election in population protocols. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA, 2018.

[39] L. Gasieniec, G. Stachowiak, and P. Uznanski. Almost logarithmic-time space optimal leader election in population protocols. *CoRR*, abs/1802.06867, 2018.

[40] G. Giakkoupis. Tight bounds for rumor spreading in graphs of a given conductance. In *Proceedings of the International Symposium on Theoretical Aspects of Computer Science (STACS)*, 2011.

[41] L. Gordon. Bounds for the distribution of the generalized variance. *The Annals of Statistics*, 17(4) :1684–1692, 1989.

[42] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, 2014.

[43] M. Harchol-Balter, T. Leighton, and D. Lewin. Resource discovery in distributed networks. In *Proceedings of the ACM Syposium on Principles of Distributed Systems (PODC)*, 1999.

[44] S. Janson. Tail bounds for sums of geometric and exponential variables. Technical report. `http://www2.math.uu.se/~svante/papers/sjN14.pdf`

[45] M. Jelasity, A. Montresor, and O. Babaoglu. Gossip-based aggregation in large dynamic networks. *ACM Trans. Comput. Syst.*, 23(3) :219–252, 2005.

[46] R. Karp, C. Schindelhauer, S. Shenker, and B. Vocking. Randomized rumor spreading. In *Proceedings of the Annual Symposium on Foundations of Computer Science (FOCS)*, 2000.

[47] D. Kempe, A. Dobra, and J. Gehrke. Gossip-based computation of aggregate information. In *Proceedings of the Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2003.

[48] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London A : Mathematical, Physical and Engineering Sciences*, 115(772) :700–721, 1927.

[49] L. Lamport, R. Shostak, and M. Pease. The byzantine generals problem. *ACM Trans. Program. Lang. Syst.*, 4(3) :382–401, July 1982.

[50] E. Lebensztayn, A. F. Machado, and P. M. Rodríguez. On the behaviour of a rumour process with random stifling. *Environmental Modelling & Software*, 26 :517–522, 2011.

[51] G. B. Mertzios, S. E. Nikoletseas, C. Raptopoulos, and P. G. Spirakis. Determining majority in networks with local interactions and very small local memory. In *Proceedings of the 41st International Colloquium (ICALP)*, pages 871–882, 2014.

[52] M. Mitzenmacher. Load balancing and density dependent jump Markov processes. In *Proceedings of International Conference on Foundations of Computer Science*, 1996.

[53] M. Mitzenmacher, A. W. Richa, and R. Sitaraman. The power of two random choices : A survey of techniques and results. In *Handbook of Randomized Computing*, pages 255–312. Kluwer, 2000.

[54] Y. Mocquard, E. Anceaume, J. Aspnes, Y. Busnel, and B. Sericola. Counting with population protocols. In *Proceedings of the 14th IEEE International Symposium on Network Computing and Applications*, pages 35–42, 2015.

[55] Y. Mocquard, E. Anceaume, and B. Sericola. Optimal Proportion Computation with Population Protocols. In *Procedings of the 15th IEEE International Symposium on Network Computing and Applications, NCA*. IEEE, 2016.

[56] Y. Mocquard, E. Anceaume, and B. Sericola. Balanced allocations and global clock in population protocols : An accurate analysis. In *SIROCCO*, June 2018.

[57] Y. Mocquard, S. Robert, B. Sericola, and E. Anceaume. Analysis of the propagation time of a rumour in large-scale distributed systems. In *Proceedings of the 15th IEEE International Symposium on Network Computing and Applications (NCA)*, 2016.

[58] Y. Mocquard, B. Sericola, and E. Anceaume. Probabilistic analysis of counting protocols in large-scale asynchronous and anonymous systems. In *16th IEEE International Symposium on Network Computing and Applications, NCA*, pages 315–322, 2017.

[59] Y. Mocquard, B. Sericola, and E. Anceaume. Population protocols with convergence detection. In *17th IEEE International Symposium on Network Computing and Applications, NCA*, 2018.

[60] Y. Mocquard, B. Sericola, and E. Anceaume. Probabilistic analysis of rumor spreading time. *INFORMS Journal on Computing*, 2018. Under press.

[61] S. Molchanov and J. M. Whitmeyer. Two Markov Models of the Spread of Rumors. *Journal of Mathematical Sociology*, 34 :157–166, 2010.

[62] O.Bournez, P.Fraigniaud, and X.Koegler. Computing with large populations using interactions. In *Mathematical Foundations of Computer Science 2012*, pages 234–246, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[63] K. Panagiotou, X. Perez-Gimenez, T. Sauerwald, and H. Sun. Randomized rumor spreading : the effect of the network topology. *Combinatorics, Probability and Computing*, 24(2) :457–479, 2015.

[64] K. Panagiotou and L. Speidel. Asynchronous rumor spreading on random graphs. *Algorithmica*, 2016.

[65] D. Peng, W. Liu, C. Lin, Z. Chen, and J. Song. A loosely synchronized gossip-based algorithm for aggregate information computation. In *2008 33rd IEEE Conference on Local Computer Networks (LCN)*, 2008.

[66] Y. Peres, K. Talwar, and U. Wieder. The $(1+\beta)$-choice process and weighted balls into bins. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2010.

[67] Y. Peres, K. Talwar, and U. Wieder. Graphical balanced allocations and the $(1 + \beta)$-choice process. *Random Structure of Algorithms*, 47(4) :760–775, 2015.

[68] E. Perron, D. Vasudevan, and M. Vojnovic. Using three states for binary consensus on complete graphs. In *Proceedings of the INFOCOM Conference*, pages 2527–2435, 2009.

[69] B. Pittel. On spreading a rumor. *SIAM Journal on Applied Mathematics*, 47(1) :213–223, 1987.

[70] M. Raab and A. Steger. "Balls into bins" — a simple and tight analysis. In *Proceedings of the Second International Workshop on Randomization and Approximation Techniques in Computer Science (RANDOM)*, pages 159–170, 1998.

[71] T. Sauerwald and H. Sun. Tight bounds for randomized load balancing on arbitrary network topologies. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 341–350, Oct 2012.

[72] B. Sericola. *Markov Chains. Theory, Algorithms and Applications.* Applied stochastic methods series. WILEY, 2013.

[73] D. W. Stroock. *Probability Theory : An Analytic View.* Cambridge University Press, second edition, 2010.

[74] K. Talwar and U. Wieder. Balanced allocations : A simple proof for the heavily loaded case. In *Automata, Languages, and Programming*, pages 979–990. Springer Berlin Heidelberg, 2014.

Annexe A

La diffusion de rumeur

À certains endroits de cette annexe, dans le but de faciliter la lecture, nous reproduisons les explications déjà données dans le corps du manuscrit.

A.1 La diffusion de rumeur en temps discret

Nous donnons, dans cette section, une démonstration des théorèmes 4.2.3, 4.2.4 et 4.2.7. Pour démontrer le théorème 4.2.3, nous avons d'abord besoin du lemme suivant.

Lemme A.1.1 *Soit* $N \geq 1$, $a \in]0;1[$, $b_1, \ldots, b_N \in]0;1[$, $c_1, \ldots, c_N \in \mathbb{R}$ *et* $d_1, \ldots, d_N \in \mathbb{R}$, *avec la condition,*

$$\textit{pour tout } j = 1, \ldots, N, \quad d_j = 0 \textit{ si } b_j = a.$$

Alors la suite $(u_k)_{k\geq 0}$ *définie par*

$$u_0 = 1 \textit{ et } u_{k+1} = au_k + \sum_{j=1}^{N} (c_j b_j + k d_j) b_j^{k-1}, \ k \geq 0 \tag{A.1}$$

vérifie

$$u_k = \left(1 - \sum_{j=1}^{N} \theta_j 1_{\{b_j \neq a\}}\right) a^k + \sum_{j=1}^{N} \left((\theta_j b_j + k\gamma_j)\, 1_{\{b_j \neq a\}} + k c_j 1_{\{b_j = a\}}\right) b_j^{k-1}, \tag{A.2}$$

avec

$$\theta_j = \frac{c_j}{b_j - a} - \frac{d_j}{(b_j - a)^2} \textit{ et } \gamma_j = \frac{d_j}{b_j - a}.$$

Preuve. Nous prouvons ce lemme par récurrence. Pour $k = 0$, la relation (A.2) donne $u_0 = 1$. Nous introduisons la notation $\alpha = \left(1 - \sum_{j=1}^{N} \theta_j 1_{\{b_j \neq a\}}\right)$ et $f_j(k) = \theta_j b_j + k\gamma_j$. La relation (A.2) peut être réécrite comme

$$u_k = \alpha a^k + \sum_{j=1}^{N} \left[f_j(k) 1_{\{b_j \neq a\}} + k c_j 1_{\{b_j = a\}}\right] b_j^{k-1}.$$

Supposons que cette dernière égalité soit vraie pour une valeur de $k \geq 0$ fixée. A partir de la relation (A.1), nous obtenons

$$\begin{aligned} u_{k+1} &= \alpha a^{k+1} + \sum_{j=1}^{N} \left[a f_j(k) 1_{\{b_j \neq a\}} + k a c_j 1_{\{b_j = a\}} \right] b_j^{k-1} + \sum_{j=1}^{N_j} (c_j b_j + k d_j) b_j^{k-1} \\ &= \alpha a^{k+1} + \sum_{j=1}^{N} \Big[a f_j(k) 1_{\{b_j \neq a\}} + k c_j b_j 1_{\{b_j = a\}} + c_j b_j + k d_j \Big] b_j^{k-1}. \end{aligned}$$

En écrivant $c_j b_j = c_j b_j 1_{\{b_j \neq a\}} + c_j b_j 1_{\{b_j = a\}}$ et $d_j = d_j 1_{\{b_j \neq a\}}$ car $d_j = 0$ quand $b_j = a$, nous obtenons

$$u_{k+1} = \alpha a^{k+1} + \sum_{j=1}^{N} \Big[(a f_j(k) + c_j b_j + k d_j) 1_{\{b_j \neq a\}} + (k+1) c_j b_j 1_{\{b_j = a\}} \Big] b_j^{k-1}. \quad \text{(A.3)}$$

Le premier terme de cette somme peut être simplifié de la manière suivante. Par définition de $f_j(k)$ et du fait que $a\gamma_j + d_j = \gamma_j b_j$, nous avons

$$\begin{aligned} a f_j(k) + c_j b_j + k d_j &= a\theta_j b_j + c_j b_j + k(a\gamma_j + d_j) \\ &= a\theta_j b_j + c_j b_j + k\gamma_j b_j \\ &= (a\theta_j + c_j + k\gamma_j) b_j \\ &= (a\theta_j + c_j - \gamma_j + (k+1)\gamma_j) b_j. \end{aligned}$$

Comme $c_j - \gamma_j = (b_j - a)\theta_j$, cette dernière expression mène à

$$a f_j(k) + c_j b_j + k d_j = (\theta_j b_j + (k+1)\gamma_j) b_j = b_j f_j(k+1).$$

En mettant ce résultat dans (A.3), nous obtenons

$$u_{k+1} = \alpha a^{k+1} + \sum_{j=1}^{N} \Big[f_j(k+1) 1_{\{b_j \neq a\}} + (k+1) c_j 1_{\{b_j = a\}} \Big] b_j^k,$$

ce qu'il fallait démontrer. ∎

Nous sommes maintenant prêts pour la preuve du théorème 4.2.3.

Théorème 4.2.3 Pour tout $n \geq 1$, $k \geq 0$ et $i \in [\![1, n-1]\!]$, nous avons

$$\mathbb{P}\{T_n > k \mid Y_0 = n - i\} = \sum_{j=1}^{\lfloor n/2 \rfloor} \left(c_{i,j}(1 - p_j) + k d_{i,j} \right) (1 - p_j)^{k-1},$$

où les coefficients $c_{i,j}$ et $d_{i,j}$, qui ne dépendent pas de k, sont donnés, pour $j \in [\![1, n-1]\!]$, par

$$c_{1,j} = 1_{\{j=1\}} \text{ et } d_{1,j} = 0$$

et pour $(i, j) \in [\![2, n-1]\!] \times [\![1, n-1]\!]$, par

$$\begin{cases} c_{i,j} &= \dfrac{p_i c_{i-1,j}}{p_i - p_j} - \dfrac{p_i d_{i-1,j}}{(p_i - p_j)^2} & \text{pour} \quad i \neq j, n-j, \\ d_{i,j} &= \dfrac{p_i d_{i-1,j}}{p_i - p_j} & \text{pour} \quad i \neq j, n-j, \\ c_{i,i} &= 1 - \displaystyle\sum_{j=1, j\neq i}^{\lfloor n/2 \rfloor} c_{i,j} & \text{pour} \quad i \leq \lfloor n/2 \rfloor, \\ c_{i,n-i} &= 1 - \displaystyle\sum_{j=1, j\neq n-i}^{\lfloor n/2 \rfloor} c_{i,j} & \text{pour} \quad i > \lfloor n/2 \rfloor, \\ d_{i,i} &= p_i c_{i-1,i} & \text{pour} \quad i \leq \lfloor n/2 \rfloor, \\ d_{i,n-i} &= p_i c_{i-1,n-i} & \text{pour} \quad i > \lfloor n/2 \rfloor. \end{cases} \tag{4.6}$$

Preuve du théorème 4.2.3. La preuve est faite par récurrence sur l'entier i. En fait, nous devons prouver que pour tout $i \in [\![1, n-1]\!]$, nous avons

$$V_{n-i}(k) = \sum_{j=1}^{n/2} \left(c_{i,j}(1-p_j) + k d_{i,j}\right)(1-p_j)^{k-1} \tag{A.4}$$

$$d_{i,j} = 0 \text{ pour } j < n-i \tag{A.5}$$

$$c_{i,j} = 0 \text{ pour } j > i. \tag{A.6}$$

Les relations (A.5) et (A.6) sont vraies pour $i = 1$ car par définition nous avons $c_{1,j} = 1_{\{j=1\}}$ et $d_{1,j} = 0$. Il s'ensuit que la relation (A.4) donne, pour $i = 1$, $V_{n-1}(k) = (1-p_1)^k$, ce qui est en accord avec la relation (4.4).

Supposons maintenant que les relations (A.4), (A.5) et (A.6) soient vraies pour un entier i fixé, $i \leq n-2$. En utilisant la relation (4.4) au point $k+1$ et le fait que $p_{n-i-1} = p_{i+1}$, nous obtenons

$$\begin{aligned} V_{n-i-1}(k+1) &= (1-p_{i+1})V_{n-i-1}(k) + p_{i+1}V_{n-i}(k) \\ &= (1-p_{i+1})V_{n-i-1}(k) + p_{i+1}\sum_{j=1}^{n/2}\left(c_{i,j}(1-p_j) + k d_{i,j}\right)(1-p_j)^{k-1}. \end{aligned} \tag{A.7}$$

L'entier i étant fixé, nous appliquons le lemme (A.1.1) en prenant comme paramètres $u_k = V_{n-i-1}(k)$, $a = 1-p_{i+1}$, $N = \lfloor n/2 \rfloor$ et, pour $j \in [\![1, N]\!]$, $b_j = 1-p_j$, $c_j = p_{i+1}c_{i,j}$ et $d_j = p_{i+1}d_{i,j}$. Notons que la condition du lemme (A.1.1) est satisfaite. En effet, $a = b_j$ est équivalent à $i+1 = j$ ou $i+1 = n-j$. Comme $i+1 = j$ est équivalent à $i+j+1 = 2j$ et comme $j \leq n/2$, ceci implique que $i+j < n$, ce qui veut dire à partir de la relation (A.5) que $d_j = p_{i+1}d_{i,j} = 0$. Avec cette notation, les paramètres θ_j et γ_j s'écrivent

$$\theta_j = \frac{c_j}{b_j - a} - \frac{d_j}{(b_j-a)^2} = \frac{p_{i+1}c_{i,j}}{p_{i+1}-p_j} - \frac{p_{i+1}d_{i,j}}{(p_{i+1}-p_j)^2},$$

$$\gamma_j = \frac{d_j}{b_j - a} = \frac{p_{i+1}d_{i,j}}{p_{i+1}-p_j}.$$

En appliquant le lemme (A.1.1) à partir de (A.7), nous obtenons

$$\begin{aligned}V_{n-i-1}(k) &= \left(1-\sum_{j=1}^{n/2}\theta_j 1_{\{i+1\neq j,n-j\}}\right)(1-p_{i+1})^k\\ &+\sum_{j=1}^{n/2}(\theta_j(1-p_j)+k\gamma_j)1_{\{i+1\neq j,n-j\}}(1-p_j)^{k-1}\\ &+\sum_{j=1}^{n/2}kp_{i+1}c_{i,j}1_{\{i+1=j \text{ or } i+1=n-j\}}(1-p_j)^{k-1}.\end{aligned}$$

En définissant

$$\begin{aligned}c_{i+1,j} &= \theta_j && \text{pour } i+1\neq j,n-j,\\ d_{i+1,j} &= \gamma_j && \text{pour } i+1\neq j,n-j,\\ c_{i+1,i+1} &= 1-\sum_{j=1}^{n/2}c_{i+1,j} && \text{pour } i+1\leq n/2,\\ c_{i+1,n-i-1} &= 1-\sum_{j=1}^{n/2}c_{i+1,j} && \text{pour } i+1> n/2,\\ d_{i+1,i+1} &= p_{i+1}c_{i,i+1} && \text{pour } i+1\leq n/2,\\ d_{i+1,n-i-1} &= p_{i+1}c_{i,n-i-1} && \text{pour } i+1> n/2,\end{aligned}$$

nous obtenons, en utilisant encore le fait que $p_{i+1}=p_{n-i-1}$,

$$\begin{aligned}V_{n-i-1}(k) &= c_{i+1,i+1}1_{\{i+1\leq n/2\}}(1-p_{i+1})^k\\ &+c_{i+1,n-i-1}1_{\{n-i-1<n/2\}}(1-p_{i+1})^k\\ &+\sum_{j=1}^{n/2}(c_{i+1,j}(1-p_j)+kd_{i+1,j})1_{\{i+1\neq j,n-j\}}(1-p_j)^{k-1}\\ &+kp_{i+1}c_{i,i+1}1_{\{i+1\leq n/2\}}(1-p_{i+1})^{k-1}\\ &+kp_{i+1}c_{i,n-i-1}1_{\{n-i-1<n/2\}}(1-p_{i+1})^{k-1},\end{aligned}$$

ce qui donne

$$V_{n-i-1}(k)=\sum_{j=1}^{n/2}\left(c_{i+1,j}(1-p_j)+kd_{i+1,j}\right)(1-p_j)^{k-1}.$$

Maintenant, nous devons prouver la récurrence pour les propriété additionnelles (A.5) et (A.6). Notons que $i+1<j$ implique que $i+1\neq j$ et $i+1\neq n-j$. Aussi, si $i+1<j$ alors nous avons $i<j$ et $i<n/2<n-j$, ce qui signifie que $c_{i,j}=d_{i,j}=0$, ce qui à son tour implique que $\theta_j=0$ et donc $c_{i+1,j}=0$

De la même manière, notons que $i+1<n-j$ implique $i+1\neq n-j$. Aussi, si $i+1<n-j$ alors nous avons $i<n-j$, ce qui signifie que $d_{i,j}=0$, ce qui à son tour implique que $\gamma_j=0$ et donc $d_{i+1,j}=0$, ceci termine la preuve. ■

La formule exacte de la distribution de T_n présentée plus haut est assez complexe à utiliser en pratique, et le calcul des coefficients du théorème 4.2.3 peut prendre beaucoup de temps pour de grandes valeurs de n. Pour simplifier ce problème, nous proposons dans le théorème suivant un majorant qui est aussi un équivalent de la quantité $\mathbb{P}\{T_n > k \mid Y_0 = i\}$. Ces calculs sont déduits de la formule de récurrence (4.4).

Théorème 4.2.4 Pour tout $n \geq 2$ et $k \geq 1$, nous avons

$$\mathbb{P}\{T_n > k \mid Y_0 = 1\} \leq \left(1 + \frac{2k(n-2)^2}{n}\right)\left(1 - \frac{2}{n}\right)^{k-1},$$

$$\mathbb{P}\{T_n > k \mid Y_0 = 1\} \underset{k\longrightarrow\infty}{\sim} \left(1 + \frac{2k(n-2)^2}{n}\right)\left(1 - \frac{2}{n}\right)^{k-1}$$

et pour $i \in [\![2, n-1]\!]$ et $k \geq 0$,

$$\mathbb{P}\{T_n > k \mid Y_0 = i\} \leq \frac{(n-i)(n-2)}{i-1}\left(1 - \frac{2}{n}\right)^{k},$$

$$\mathbb{P}\{T_n > k \mid Y_0 = i\} \underset{k\longrightarrow\infty}{\sim} \frac{(n-i)(n-2)}{i-1}\left(1 - \frac{2}{n}\right)^{k}.$$

D'autre part, nous avons

$$\mathbb{P}\{T_n > k\} \leq \mathbb{P}\{T_n > k \mid Y_0 = 1\}.$$

Preuve du théorème 4.2.4. Le résultat est trivial pour $n = 2$ car nous avons, dans ce cas, $T_2 = 1$. Nous supposons donc que $n \geq 3$. Notons que par définition de p_i nous avons $p_i = p_{n-i}$. Considérons la suite b_i définie pour $i \in \{1, \ldots, n-2\}$, par

$$b_1 = 1 \text{ et } b_i = \frac{p_i b_{i-1}}{p_i - p_1}, \text{ pour } i \in [\![2, n-2]\!].$$

En observant que

$$b_i = \frac{i(n-i)b_{i-1}}{(i-1)(n-i-1)},$$

on vérifie facilement par récurrence que pour $i \in [\![1, n-2]\!]$, nous avons

$$b_i = \frac{i(n-2)}{n-i-1}.$$

Nous allons montrer maintenant par récurrence que pour tout $i \in [\![1, n-2]\!]$, nous avons

$$V_{n-i}(k) \leq b_i\,(1-p_1)^k, \text{ pour tout } k \geq 0$$
$$\text{et } V_{n-i}(k) \underset{k\longrightarrow\infty}{\sim} b_i\,(1-p_1)^k.$$

Ces deux résultats sont vrais pour $i = 1$ car $V_{n-1}(k) = (1 - p_{n-1})^k = (1 - p_1)^k$. Supposons maintenant que ce résultat soit vrai pour un entier i fixé avec $1 \leq i \leq n - 3$. A partir des égalités (4.4), nous avons

$$\begin{aligned} V_{n-i-1}(k) &= (1 - p_{n-i-1})V_{n-i-1}(k-1) + p_{n-i-1}V_{n-i}(k-1) \\ &= (1 - p_{i+1})V_{n-i-1}(k-1) + p_{i+1}V_{n-i}(k-1). \end{aligned}$$

En utilisant l'hypothèse de récurrence, nous obtenons, en ce qui concerne l'inégalité

$$V_{n-i-1}(k) \leq (1 - p_{i+1})V_{n-i-1}(k-1) + p_{i+1}b_i \left(1 - p_1\right)^{k-1}.$$

En développant cette inégalité et en utilisant le fait que $V_{n-i-1}(0) = 1$, nous arrivons à

$$\begin{aligned} V_{n-i-1}(k) &\leq (1 - p_{i+1})^k + p_{i+1}b_i \sum_{j=0}^{k-1} (1 - p_{i+1})^j (1 - p_1)^{k-1-j} \\ &= (1 - p_{i+1})^k + p_{i+1}b_i \frac{(1 - p_1)^k - (1 - p_{i+1})^k}{p_{i+1} - p_1} \\ &= (1 - p_{i+1})^k + b_{i+1}\left((1 - p_1)^k - (1 - p_{i+1})^k\right) \\ &= (1 - b_{i+1})\left(1 - p_{i+1}\right)^k + b_{i+1}(1 - p_1)^k. \end{aligned}$$

Comme $b_{i+1} \geq 1$, nous avons

$$V_{n-i-1}(k) \leq b_{i+1}(1 - p_1)^k.$$

En utilisant un calcul similaire, nous obtenons

$$V_{n-i-1}(k) \underset{k \longrightarrow \infty}{\sim} (1 - b_{i+1})\left(1 - p_{i+1}\right)^k + b_{i+1}(1 - p_1)^k.$$

Comme $p_{i+1} > p_1$, nous avons aussi

$$V_{n-i-1}(k) \underset{k \longrightarrow \infty}{\sim} b_{i+1}(1 - p_1)^k.$$

Ainsi nous avons montré que pour tout $i \in [\![1, n-2]\!]$, nous avons

$$\begin{aligned} V_{n-i}(k) &\leq b_i \left(1 - p_1\right)^k, \text{ for all } k \geq 0 \\ \text{et } V_{n-i}(k) &\underset{k \longrightarrow \infty}{\sim} b_i \left(1 - p_1\right)^k. \end{aligned}$$

En particulier, pour $i = n - 2$ nous obtenons

$$\begin{aligned} V_2(k) &\leq b_{n-2} \left(1 - p_1\right)^k, \text{ pour tout } k \geq 0 \\ \text{et } V_2(k) &\underset{k \longrightarrow \infty}{\sim} b_{n-2} \left(1 - p_1\right)^k. \end{aligned}$$

Considérons maintenant le terme $V_1(k)$. A partir des relations (4.4) et utilisant l'inégalité précédente, nous avons

$$\begin{aligned} V_1(k) &= (1 - p_1)V_1(k-1) + p_1V_2(k-1) \\ &\leq (1 - p_1)V_1(k-1) + p_1b_{n-2} \left(1 - p_1\right)^{k-1}. \end{aligned}$$

En développant cette inégalité et en utilisant le fait que $V_1(0) = 1$, cela mène à

$$\begin{aligned} V_1(k) &\leq (1-p_1)^k + p_1 b_{n-2} \sum_{j=0}^{k-1} (1-p_1)^j (1-p_1)^{k-1-j} \\ &= (1-p_1)^k + p_1 b_{n-2} k (1-p_1)^{k-1} \\ &= (1 - p_1 + k p_1 b_{n-2}) (1-p_1)^{k-1} \\ &\leq (1 + k p_1 b_{n-2}) (1-p_1)^{k-1}, \end{aligned}$$

ce qui donne

$$V_1(k) \leq \left(1 + \frac{2k(n-2)^2}{n}\right)\left(1 - \frac{2}{n}\right)^{k-1}.$$

De manière similaire, nous obtenons

$$V_1(k) \underset{k \longrightarrow \infty}{\sim} \left(1 + \frac{2k(n-2)^2}{n}\right)\left(1 - \frac{2}{n}\right)^{k-1}.$$

Finalement, comme $\mathbb{P}\{T_n > k \mid Y_0 = i\}$ est décroissante en fonction de i, nous avons

$$\begin{aligned} \mathbb{P}\{T_n > k\} &= \sum_{i=1}^{n-1} \mathbb{P}\{T_n > k \mid Y_0 = i\} \mathbb{P}\{Y_0 = i\} \\ &\leq \mathbb{P}\{T_n > k \mid Y_0 = 1\}, \end{aligned}$$

ce qui termine la preuve. ■

En ce qui concerne le théorème suivant, rappelons que $f(c,n)$ est issue du théorème précédent tandis que $g(c,n)$ est issue du théorème 4.2.6 qui utilise la borne de Chernoff. La valeur c^* représente la valeur à partir de laquelle $f(c,n)$ est un meilleur majorant que $g(c,n)$. Voici les formules explicites des ces fonctions

$$\begin{aligned} f(c,n) &= \left(1 + \frac{2c(n-1)H_{n-1}(n-2)^2}{n}\right)\left(1 - \frac{2}{n}\right)^{c(n-1)H_{n-1}-2} \\ g(c,n) &= \frac{1}{c}\left(1 - \frac{2}{n}\right)^{(c-1-\ln(c))(n-1)H_{n-1}} \\ e(c,n) &= \mathbb{P}\{T_n > c\mathbb{E}(T_n)\}. \end{aligned}$$

Théorème 4.2.7 Pour tout $n \geq 3$, il existe un unique $c^* \geq 1$ tel que $f(c^*,n) = g(c^*,n)$ et nous avons

$$\begin{cases} f(c,n) > g(c,n) & \text{pour tout } 1 \leq c < c^* \\ \\ f(c,n) < g(c,n) & \text{pour tout } c > c^*. \end{cases} \tag{4.7}$$

De plus,

$$\lim_{c \longrightarrow \infty} \frac{f(c,n)}{g(c,n)} = 0.$$

Preuve du théorème 4.2.7. Pour simplifier l'écriture, nous introduisons la notation suivante.

$$\begin{aligned} A_n =& (n-1)H_{n-1}\ln\left(1-\frac{2}{n}\right) \\ B_n =& \frac{2(n-1)H_{n-1}(n-2)^2}{n} \\ C_n =& \left(1-\frac{2}{n}\right)^{(n-1)H_{n-1}-2}. \end{aligned}$$

Premièrement, comme

$$\left(1-\frac{2}{n}\right)^{\ln(c)(n-1)H_{n-1}} = c^{(n-1)H_{n-1}\ln(1-2/n)} = c^{A_n},$$

nous avons

$$\frac{f(c,n)}{g(c,n)} = C_n(1+cB_n)c^{A_n+1}.$$

En prenant la dérivée par rapport à c, nous obtenons

$$\frac{d(f/g)}{dc}(c,n) = C_n c^{A_n}[A_n+1+cB_n(A_n+2)].$$

Le terme $C_n c^{A_n}$ est strictement positif pour tout $c \geq 1$, aussi nous avons

$$\frac{d(f/g)}{dc}(c,n) \geq 0 \Longleftrightarrow c \leq -\frac{A_n+1}{B_n(A_n+2)},$$

On vérifie facilement que $-(A_n+1)/(B_n(A_n+2)) < 0$ pour tout $n \geq 3$, ce qui signifie que $d(f/g)/dc < 0$ pour tout $c \geq 1$. Ce qui implique que la fonction $c \longmapsto f(c,n)/g(c,n)$ est strictement décroissante sur $[1,+\infty[$.

En observant que $A_n < -2$ pour tout $n \geq 3$, nous obtenons

$$\lim_{c\longrightarrow\infty}\frac{f(c,n)}{g(c,n)} = 0,$$

ce qui prouve la deuxième partie du théorème. Deuxièmement comme

$$\frac{f(1,n)}{g(1,n)} = C_n(1+B_n),$$

on vérifie facilement que $f(1,n)/g(1,n) \geq 1$, pour tout $n \geq 3$.

Récapitulons et concluons. Pour tout $n \geq 3$, nous avons $f(1,n)/g(1,n) \geq 1$ et $c \longmapsto f(c,n)/g(c,n)$ est continue, strictement décroissante et tend vers 0. Il s'ensuit qu'il existe une unique valeur $c^* \geq 1$ telle que $f(c^*,n)/g(c^*,n) = 1$ et satisfaisant les conditions (4.7), ce qui termine la preuve. ■

Pour prouver le théorème 4.2.10, nous avons d'abord besoin d'un lemme technique.

Les variables aléatoires S_i sont indépendantes et distribuées selon la loi géométrique de paramètre $p_i = 2i(n-i)/(n(n-1))$, mais, comme les p_i dépendent de n, nous renommons les variables aléatoires S_i en $S_{n,i}$ et les paramètres p_i en $p_{n,i}$.

Le temps de diffusion T_n s'écrit donc $T_n = S_{n,1} + \cdots + S_{n,n-1}$.

Nous utilisons la notation $X_n \xrightarrow{\mathcal{L}} X$ pour exprimer le fait que la suite de variables aléatoires (X_n) converge en distribution (ou en loi) vers la variable aléatoire X quand n tend vers l'infini, s'il ne s'agit pas de n, c'est précisé. De même nous utilisons la notation $X \overset{\mathcal{L}}{=} Y$ pour exprimer que X suit la même loi que Y.

Lemme A.1.2 *Soit $(Z_i)_{i\geq 1}$ une suite de variables aléatoires i.i.d., de loi exponentielle de paramètre 1 et soit W défini par*

$$W = \sum_{i=1}^{\infty} \frac{Z_i - 1}{2i}.$$

Alors nous avons

$$\frac{T_n - \mathbb{E}(T_n)}{n} \xrightarrow{\mathcal{L}} W^{(1)} + W^{(2)}$$

où $W^{(1)}$ et $W^{(2)}$ sont i.i.d. de même loi que W.

Preuve. Pour tout i, nous avons $\lim_{n\longrightarrow\infty} p_{n,i} = 0$. Il s'ensuit que pour tout $x \geq 0$, nous avons

$$\mathbb{P}\{p_{n,i}S_{n,i} > x\} = \mathbb{P}\{S_{n,i} > x/p_{n,i}\} = (1-p_{n,i})^{\lfloor x/p_{n,i}\rfloor} \text{ et } \lim_{n\longrightarrow\infty}(1-p_{n,i})^{\lfloor x/p_{n,i}\rfloor} = e^{-x}.$$

Si Z_i est une variable aléatoire de loi exponentielle de paramètre 1, nous avons montré que $p_{n,i}S_{n,i} \xrightarrow{\mathcal{L}} Z_i$. De plus, comme les $(S_{n,i})_{i\in[\![1,n-1]\!]}$ sont indépendantes, les $(Z_i)_{i\geq 1}$ le sont aussi.

En observant maintenant que pour chaque i, nous avons $\lim_{n\longrightarrow\infty} np_{n,i} = 2i$ et en définissant $R_{n,i} = S_{n,i} - \mathbb{E}(S_{n,i})$, nous obtenons, du fait que $\mathbb{E}(S_{n,i}) = 1/p_{n,i}$,

$$\frac{R_{n,i}}{n} = \frac{S_{n,i} - \mathbb{E}(S_{n,i})}{n} = \frac{p_{n,i}S_{n,i} - 1}{np_{n,i}} \xrightarrow{\mathcal{L}} \frac{Z_i - 1}{2i}. \tag{A.8}$$

Supposons maintenant que $n = 2k+1$. Alors nous avons

$$\frac{T_{2k+1} - \mathbb{E}(T_{2k+1})}{2k+1} = \frac{1}{2k+1}\left(\sum_{i=1}^{k} R_{2k+1,i} + \sum_{i=1}^{k} R_{2k+1,2k+1-i}\right) = V_k + \overline{V}_k, \tag{A.9}$$

où

$$V_k = \frac{1}{2k+1}\sum_{i=1}^{k} R_{2k+1,i} \text{ et } \overline{V}_k = \frac{1}{2k+1}\sum_{i=1}^{k} R_{2k+1,2k+1-i}.$$

Les variables aléatoires V_k et $\overline{V}_k$ sont indépendantes et ont aussi la même distribution. En effet, comme $p_{n,i} = p_{n,n-i}$ les variables $R_{n,i}$ et $R_{n,n-i}$ ont la même distribution.

Le reste de la preuve consiste à vérifier les hypothèses du principe des lois accompagnantes (accompanying laws) du théorème 3.1.14 de [73]. Nous introduisons la notation

$$W_{m,k} = \frac{1}{2k+1} \sum_{i=1}^{m-1} R_{2k+1,i}.$$

En utilisant le fait que $\mathbb{E}(R_{n,i}) = 0$ et que les $R_{n,i}$ sont indépendantes, nous avons

$$\begin{aligned}
\mathbb{E}((V_k - W_{m,k})^2) &= \mathbb{E}\left(\left[\frac{1}{2k+1}\sum_{i=m}^{k} R_{2k+1,i}\right]^2\right) = \mathbb{V}\left(\frac{1}{2k+1}\sum_{i=m}^{k} R_{2k+1,i}\right) \\
&= \frac{1}{(2k+1)^2}\sum_{i=m}^{k} \mathbb{V}(R_{2k+1,i}) = \frac{1}{(2k+1)^2}\sum_{i=m}^{k} \mathbb{V}(S_{2k+1,i}) \\
&= \frac{1}{(2k+1)^2}\sum_{i=m}^{k} \frac{1-p_{2k+1,i}}{p_{2k+1,i}^2} \leq \frac{1}{(2k+1)^2}\sum_{i=m}^{k} \frac{1}{p_{2k+1,i}^2}.
\end{aligned}$$

En rappelant que $p_{2k+1,i} = 2i(2k+1-i)/(2k(2k+1))$, nous obtenons

$$\mathbb{E}((V_k - W_{m,k})^2) \leq k^2 \sum_{i=m}^{k} \frac{1}{i^2(2k+1-i)^2}.$$

Dans cette somme nous avons $2k+1-i \geq k$. Cela mène à

$$\mathbb{E}((V_k - W_{m,k})^2) \leq \sum_{i=m}^{k} \frac{1}{i^2}.$$

Alors nous avons

$$\lim_{m\longrightarrow\infty} \limsup_{k\longrightarrow\infty} \mathbb{E}((V_k - W_{m,k})^2) \leq \lim_{m\longrightarrow\infty} \sum_{i=m}^{\infty} \frac{1}{i^2} = 0.$$

D'un autre coté en utilisant l'inégalité de Markov, nous obtenons, pour tout $\varepsilon > 0$,

$$\mathbb{P}\{|V_k - W_{m,k}| \geq \varepsilon\} = \mathbb{P}\{(V_k - W_{m,k})^2 \geq \varepsilon^2\} \leq \frac{\mathbb{E}((V_k - W_{m,k})^2)}{\varepsilon^2}.$$

En mettant ensemble ces résultats, nous avons montré que pour tout $\varepsilon > 0$, nous avons

$$\lim_{m\longrightarrow\infty} \limsup_{k\longrightarrow\infty} \mathbb{P}\{|V_k - W_{m,k}| \geq \varepsilon\} = 0. \tag{A.10}$$

Nous introduisons la notation

$$W_m = \sum_{i=1}^{m-1} \frac{Z_i - 1}{2i}.$$

En utilisant (A.8) et le fait que les $R_{n,i}$ sont indépendants, nous avons

$$W_{m,k} \xrightarrow{\mathcal{L}} W_m \text{ quand } k \longrightarrow \infty. \tag{A.11}$$

Les hypothèses du principe des lois accompagnantes (accompanying laws) du théorème 3.1.14 de [73] sont les propriétés (A.8) et (A.11). Par conséquent, nous pouvons conclure que

$$V_k \xrightarrow{\mathcal{L}} W \text{ quand } k \longrightarrow \infty.$$

De manière similaire, nous avons

$$\overline{V}_k \xrightarrow{\mathcal{L}} W \text{ quand } k \longrightarrow \infty.$$

Ce qui signifie, en utilisant la relation (A.9), que

$$\frac{T_{2k+1} - \mathbb{E}(T_{2k+1})}{2k+1} \xrightarrow{\mathcal{L}} W^{(1)} + W^{(2)},$$

où $W^{(1)}$ et $W^{(2)}$ sont i.i.d. et de même loi que W. Le même raisonnement s'applique de manière identique dans le cas où $n = 2k$. ■

Nous somme maintenant prêts à démontrer le théorème 4.2.10.

Théorème 4.2.10

$$\lim_{n\longrightarrow\infty} \mathbb{P}\{T_n > \mathbb{E}(T_n)\} = 1 - 2e^{-\gamma} K_1\left(2e^{-\gamma}\right) \approx 0.448429663727.$$

où γ est la constante d'Euler donnée par $\gamma = \lim_{n\longrightarrow\infty}(H_n - \ln(n)) \approx 0.5772156649$ et K_1 est la fonction de Bessel modifiée de deuxième espèce et d'ordre 1, donnée, pour $z > 0$, par

$$K_1(z) = \frac{z}{4}\int_0^{+\infty} t^{-2} e^{-t - z^2/4t} dt.$$

Preuve du théorème 4.2.10. Louis Gordon a prouvé dans [41] que

$$-\gamma + \sum_{i=1}^{+\infty} \frac{1 - Z_i}{i} \overset{\mathcal{L}}{=} \ln Z_1,$$

où (Z_i) sont i.i.d. de loi exponentielle de paramètre 1 et γ est la constante d'Euler. Par conséquent, par définition de W dans le lemme A.1.2, nous avons

$$W \overset{\mathcal{L}}{=} -\frac{\gamma + \ln Z_1}{2}.$$

En introduisant $W^{(1)} \overset{\mathcal{L}}{=} -(\gamma + \ln Z_1)/2$ et $W^{(2)} \overset{\mathcal{L}}{=} -(\gamma + \ln Z_2)/2$, nous obtenons à partir du lemme A.1.2,

$$\begin{aligned}
\lim_{n\longrightarrow\infty} \mathbb{P}\left\{T_n > \mathbb{E}(T_n)\right\} &= \mathbb{P}\left\{W^{(1)} + W^{(2)} > 0\right\} \\
&= \mathbb{P}\left\{-2\gamma - \ln\left(Z_1 Z_2\right) > 0\right\} \\
&= \mathbb{P}\left\{Z_1 Z_2 < e^{-2\gamma}\right\} \\
&= \int_0^{\infty} \left(1 - \exp(-e^{-2\gamma}/t)\right) e^{-t} dt \\
&= 1 - \int_0^{\infty} \exp(-t - e^{-2\gamma}/t) dt.
\end{aligned}$$

Soit u la fonction définie sur $]0,+\infty[$ par $u(t) = \exp(-t - e^{-2\gamma}/t)$. Nous obtenons facilement

$$\lim_{t\longrightarrow 0^+} u(t) = 0 \text{ et } \lim_{t\longrightarrow\infty} u(t) = 0,$$

ce qui implique que

$$\int_0^\infty u'(t)dt = 0. \tag{A.12}$$

La dérivée u' de u est donnée par

$$\begin{aligned} u'(t) &= \left(-1 + e^{-2\gamma}t^{-2}\right) u(t) \\ &= -u(t) + e^{-2\gamma}u(t)t^{-2}. \end{aligned} \tag{A.13}$$

En intégrant (A.13) sur $]0,+\infty[$ et en utilisant (A.12), nous obtenons

$$\int_0^\infty u(t)dt = e^{-2\gamma}\int_0^\infty u(t)t^{-2}dt.$$

Par définition de la fonction u, cela mène à

$$\begin{aligned} \lim_{n\longrightarrow\infty} \mathbb{P}\left\{T_n > \mathbb{E}(T_n)\right\} &= 1 - e^{-2\gamma}\int_0^\infty t^{-2}\exp(-t - e^{-2\gamma}/t)dt \\ &= 1 - 2e^{-\gamma}K_1\left(2e^{-\gamma}\right) \approx 0.448429663727, \end{aligned}$$

où K_1 est la fonction bien connue de Bessel modifiée de deuxième espèce et d'ordre 1, voir par exemple l'expression 8.432.6 de [42]. ■

A.2 La diffusion de rumeur en temps continu

Pour prouver le théorème 4.3.8, nous avons d'abord besoin, comme dans le cas du temps discret, du lemme technique suivant. Rappelons que le temps de diffusion Θ_n peut s'écrire $\Theta_n = U_{n,1} + \cdots + U_{n,n-1}$ où $U_{n,i}$ suit la loi exponentielle de paramètre $\mu_{n,i} = n\lambda p_i$. Notons que le résultat suivant a aussi été obtenu par Stanislav Molchanov et Joseph M. Whitmeyer [61] en utilisant une démonstration différente.

Lemme A.2.1 *Soit $(Z_i)_{i\geq 1}$ une suite de variables aléatoires i.i.d. de loi exponentielle, de paramètre 1 et soit W définie par*

$$W = \sum_{i=1}^{\infty} \frac{Z_i - 1}{2\lambda i}.$$

Alors nous avons

$$\Theta_n - \mathbb{E}(\Theta_n) \xrightarrow{\mathcal{L}} W^{(1)} + W^{(2)}$$

où $W^{(1)}$ et $W^{(2)}$ sont i.i.d. de même loi que W.

Preuve. Pour tout $n \geq 2$, $i \in [\![1, n-1]\!]$ et $x \geq 0$, nous avons

$$\mathbb{P}\{\mu_{n,i}U_{n,i} > x\} = \mathbb{P}\{U_{n,i} > x/\mu_{n,i}\} = e^{-x}.$$

Par conséquent si Z_i est une variable aléatoire de loi exponentielle de paramètre 1, nous avons $\mu_{n,i}U_{n,i} \stackrel{\mathcal{L}}{=} Z_i$. De plus, comme les $(U_{n,i})_{i\in\{1,\ldots,n-1\}}$ sont indépendantes, les $(Z_i)_{i\geq 1}$ sont aussi indépendantes.

En observant maintenant que pour chaque i, nous avons $\lim_{n\longrightarrow\infty} \mu_{n,i} = 2\lambda i$ et en définissant $R_{n,i} = U_{n,i} - \mathbb{E}(U_{n,i})$, du fait que $\mathbb{E}(U_{n,i}) = 1/\mu_{n,i}$, nous obtenons

$$R_{n,i} = U_{n,i} - \mathbb{E}(U_{n,i}) = \frac{\mu_{n,i}S_{n,i} - 1}{\mu_{n,i}} \xrightarrow{\mathcal{L}} \frac{Z_i - 1}{2\lambda i}. \tag{A.14}$$

Supposons que $n = 2k + 1$ et définissons

$$V_k = \sum_{i=1}^{k} R_{2k+1,i} \text{ et } \overline{V}_k = \sum_{i=1}^{k} R_{2k+1,2k+1-i},$$

nous avons

$$\Theta_{2k+1} - \mathbb{E}(\Theta_{2k+1}) = V_k + \overline{V}_k. \tag{A.15}$$

Les variables aléatoires V_k et $\overline{V}_k$ sont indépendantes et ont la même loi. En effet, comme $\mu_{n,i} = \mu_{n,n-i}$, les variables $R_{n,i}$ et $R_{n,n-i}$ ont la même distribution.

Comme dans le cas du temps discret, le reste de la preuve consiste à vérifier les hypothèses du principe des lois accompagnantes du théorème 3.1.14 de [73]. Nous introduisons la notation

$$W_{m,k} = \sum_{i=1}^{m-1} R_{2k+1,i}.$$

En utilisant le fait que $\mathbb{E}(R_{n,i}) = 0$ et que les $R_{n,i}$ sont indépendantes, nous avons

$$\begin{aligned}\mathbb{E}((V_k - W_{m,k})^2) &= \mathbb{E}\left(\left[\sum_{i=m}^{k} R_{2k+1,i}\right]^2\right) = \mathbb{V}\left(\sum_{i=m}^{k} R_{2k+1,i}\right) \\ &= \sum_{i=m}^{k} \mathbb{V}(R_{2k+1,i}) = \sum_{i=m}^{k} \mathbb{V}(U_{2k+1,i}) = \sum_{i=m}^{k} \frac{1}{\mu_{2k+1,i}^2}.\end{aligned}$$

En rappelant que $\mu_{2k+1,i} = 2\lambda i(2k + 1 - i)/(2k)$, nous obtenons

$$\mathbb{E}((V_k - W_{m,k})^2) = \frac{k^2}{\lambda^2} \sum_{i=m}^{k} \frac{1}{i^2(2k + 1 - i)^2}.$$

Dans cette somme nous avons $2k + 1 - i \geq k$. Cela mène à

$$\mathbb{E}((V_k - W_{m,k})^2) \leq \frac{1}{\lambda^2} \sum_{i=m}^{k} \frac{1}{i^2}.$$

Nous obtenons

$$\lim_{m\longrightarrow\infty}\limsup_{k\longrightarrow\infty}\mathbb{E}((V_k-W_{m,k})^2)\leq\frac{1}{\lambda^2}\lim_{m\longrightarrow\infty}\sum_{i=m}^{\infty}\frac{1}{i^2}=0.$$

En introduisant la variable aléatoire

$$W_m=\sum_{i=1}^{m-1}\frac{Z_i-1}{2\lambda i},$$

le reste de la preuve est exactement le même que dans le cas du temps discret. ■

Nous sommes maintenant prêts à démontrer le théorème 4.3.8.

Théorème 4.3.8

$$\lim_{n\longrightarrow\infty}\mathbb{P}\{\Theta_n>\mathbb{E}(\Theta_n)\}=1-2e^{-\gamma}K_1\left(2e^{-\gamma}\right)\approx 0.448429663727.$$

où γ est la constante d'Euler donnée par $\gamma=\lim_{n\longrightarrow\infty}(H_n-\ln(n))\approx 0.5772156649$ et K_1 est la fonction de Bessel modifiée de deuxième espèce et d'ordre 1, donnée pour $z>0$, par

$$K_1(z)=\frac{z}{4}\int_0^{+\infty}t^{-2}e^{-t-z^2/4t}dt.$$

Preuve du théorème 4.3.8 Louis Gordon a prouvé dans [41] que

$$-\gamma+\sum_{i=1}^{+\infty}\frac{1-Z_i}{i}\overset{\mathcal{L}}{=}\ln Z_1,$$

où (Z_i) sont i.i.d. de loi exponentielle de paramètre 1 et γ est la constante d'Euler-Mascheroni. Donc, par définition de W dans le lemme A.2.1, nous avons

$$W\overset{\mathcal{L}}{=}-\frac{\gamma+\ln Z_1}{2\lambda}.$$

En introduisant $W^{(1)}\overset{\mathcal{L}}{=}-(\gamma+\ln Z_1)/2\lambda$ et $W^{(2)}\overset{\mathcal{L}}{=}-(\gamma+\ln Z_2)/2\lambda$, nous obtenons à partir du lemme A.2.1,

$$\lim_{n\longrightarrow\infty}\mathbb{P}\left\{T_n>\mathbb{E}(T_n)\right\}=\mathbb{P}\left\{W^{(1)}+W^{(2)}>0\right\}=\mathbb{P}\left\{-2\gamma-\ln\left(Z_1Z_2\right)>0\right\}.$$

Le reste de la preuve est similaire à celle du théorème 4.2.10. ■

Annexe B

Protocoles basé sur la moyenne

À certains endroits de cette annexe, dans le but de faciliter la lecture, nous reproduisons les explications déjà données dans le corps du manuscrit.

B.1 Moyenne avec des entiers

Nous allons d'abord traiter un cas particulier où la moyenne de la somme a une partie fractionnaire égale à 0.5. Dans ce cas, et dans ce cas seulement, nous pouvons appliquer l'inégalité de Markov, et obtenir une formule avec des coefficients assez faibles.

Théorème 5.2.8 Pour tout $\delta \in]0;1[$, si $\ell - \lfloor \ell \rfloor = 1/2$ et s'il existe une constante K telle que $\|C_0 - L\|_\infty \leq K$, alors, pour tout $t \geq (n-1)\,(2\ln K + \ln n - \ln \delta - \ln 2)$, nous avons

$$\mathbb{P}\{\|C_t - L\|_\infty \neq 1/2\} \leq \delta.$$

Preuve du théorème 5.2.8 Si $\ell - \lfloor \ell \rfloor = 1/2$, alors, comme pour tout $i \in [\![1, n]\!]$, $C_t^{(i)}$ est un entier, nous avons $|C_t^{(i)} - \ell| \geq 1/2$. Il s'ensuit que

$$\|C_t - L\|^2 \geq \frac{n}{4}.$$

S'il existe i tel que $|C_t^{(i)} - \ell| > 1/2$ alors nous avons $\|C_t - L\|^2 > n/4$. Inversement, si pour tout i, $|C_t^{(i)} - \ell| = 1/2$, alors nous avons $\|C_t - L\|^2 = n/4$. Nous avons donc montré que

$$\|C_t - L\|^2 = \frac{n}{4} \Longleftrightarrow \|C_t - L\|_\infty = \frac{1}{2}. \tag{B.1}$$

Par conséquent, si $\|C_t - L\|^2 > n/4$ alors il existe i tel que $|C_t^{(i)} - \ell| > 1/2$. Dans ce cas et pour cette valeur de i, comme les $C_t^{(j)}$ sont des entiers et comme $\ell - \lfloor \ell \rfloor = 1/2$, nous avons nécessairement $|C_t^{(i)} - \ell| \geq 3/2$. Donc dans ce cas nous avons

$$\|C_t - L\|^2 \geq (n-1)\left(\frac{1}{2}\right)^2 + \left(\frac{3}{2}\right)^2 = \frac{n}{4} + 2,$$

ce qui nous permet d'écrire

$$\|C_t - L\|^2 > \frac{n}{4} \Longleftrightarrow \|C_t - L\|^2 - \frac{n}{4} \geq 2. \tag{B.2}$$

A partir du théorème 5.2.4 nous obtenons, pour tout $t \geq 0$,

$$\mathbb{E}(\|C_t - L\|^2 - n/4) \leq \left(1 - \frac{1}{n-1}\right)^t \mathbb{E}(\|C_0 - L\|^2).$$

Soit $\tau = (n-1)\,(2\ln K + \ln n - \ln \delta - \ln 2)$. Pour $t \geq \tau$, nous avons

$$\left(1 - \frac{1}{n-1}\right)^t \leq e^{-t/(n-1)} \leq e^{-\tau/(n-1)} = \frac{2\delta}{nK^2}.$$

De plus , comme $\|C_0 - L\|^2 \leq n\|C_0 - L\|_\infty^2 \leq nK^2$, nous obtenons $\mathbb{E}(\|C_0 - L\|^2) \leq nK^2$ et par conséquent $\mathbb{E}(\|C_t - L\|^2 - n/4) \leq 2\delta$. En utilisant l'inégalité de Markov, pour $t \geq \tau$, nous obtenons

$$\mathbb{P}\{\|C_t - L\|^2 - n/4 \geq 2\} \leq \delta.$$

En considérant ensemble les deux équivalences (B.1) et (B.2), et d'autre part puisque $\|C_t - L\|^2 \geq n/4$, nous avons

$$\|C_t - L\|^2 - \frac{n}{4} < 2 \Longleftrightarrow \|C_t - L\|_\infty = \frac{1}{2}$$

et alors, pour $t \geq \tau$,

$$\mathbb{P}\{\|C_t - L\|_\infty \neq 1/2\} = \mathbb{P}\{\|C_t - L\|^2 - n/4 \geq 2\} \leq \delta,$$

ce qu'il fallait démontrer. ■

Voici un théorème qui donne une formule explicite du temps nécessaire pour que le vecteur C_t se rapproche du vecteur L au point d'appartenir à la boule de rayon $\sqrt{\rho n}$ et de centre L, ceci avec une probabilité supérieure à $1 - \delta$, pour $\delta \in]0;1[$ et $\rho > 1/4$. μ est un paramètre du théorème que nous fixerons ultérieurement afin d'optimiser les coefficients.

Théorème 5.2.9 Soient $\delta \in]0;1[$, $\rho > 1/4$ et $\mu > \max\{4/(4\rho - 1), 4(e-1)\}$. S'il existe $K \geq \max\left\{\sqrt{\rho n}, \|C_0 - L\|\right\}$, alors pour tout $t \geq n\theta$, nous avons

$$\mathbb{P}\left\{\|C_t - L\|^2 < \rho n\right\} \geq 1 - \delta,$$

où

$$\theta = 2\ln K - \ln n + \ln \mu - (\ln(\rho\mu)/[\ln(4\rho\mu) - \ln(\mu + 4)])\ln \delta.$$

Preuve du théorème 5.2.9. Pour tout $t \geq 0$ nous notons $Y_t = \|C_t - L\|^2$.

Soit $(T_k)_{k\geq 0}$ la suite d'instants définie par $T_0 = 0$ et

$$T_{k+1} = T_k + \left\lceil (n-1)\ln\left(\frac{\mu \mathbb{E}(Y_{T_k} \mid Y_{T_k} \geq \rho n)}{n}\right)\right\rceil. \tag{B.3}$$

En utilisant le théorème 5.2.5, nous avons, pour tout $k \geq 0$, en prenant $y = \rho n$, $t = T_{k+1}$ et $s = T_k$,

$$\mathbb{E}\left(Y_{T_{k+1}} \mid Y_{T_k} \geq \rho n\right) \leq \left(1 - \frac{1}{n-1}\right)^{T_{k+1}-T_k} \mathbb{E}\left(Y_{T_k} \mid Y_{Tk} \geq \rho n\right) + \frac{n}{4}.$$

Puisque pour tout $x \in [0, 1[$, $1 - x \leq e^{-x}$ et par définition de la suite (T_k), nous avons

$$\left(1 - \frac{1}{n-1}\right)^{T_{k+1}-T_k} \leq e^{-(T_{k+1}-T_k)/(n-1)} \leq \frac{n}{\mu \mathbb{E}\left(Y_{T_k} \mid Y_{T_k} \geq \rho n\right)}.$$

Cela mène à

$$\mathbb{E}\left(Y_{T_{k+1}} \mid Y_{T_k} \geq \rho n\right) \leq \frac{n}{\mu} + \frac{n}{4} = \frac{(\mu+4)n}{4\mu}. \tag{B.4}$$

En utilisant l'inégalité de Markov, nous obtenons

$$\mathbb{P}\{Y_{T_{k+1}} \geq \rho n \mid Y_{T_k} \geq \rho n\} \leq \frac{\mathbb{E}\left(\rho Y_{T_{k+1}} \mid Y_{T_k} \geq \rho n\right)}{\rho n} \leq \frac{\mu+4}{4\rho\mu}.$$

Nous pouvons noter au passage que, de par la définition de μ, et c'est la clef de cette démonstration, nous avons

$$4\rho\mu > \mu + 4 \iff \frac{\mu+4}{4\rho\mu} < 1 \tag{B.5}$$

Nous introduisons la suite $(\alpha_k)_{k\geq 0}$ définie par

$$\alpha_0 = \frac{(\mu+4)n}{4\mu K^2} \text{ et } \alpha_k = \max\left\{\mathbb{P}\{Y_{T_k} \geq \rho n \mid Y_{T_{k-1}} \geq \rho n\}, \alpha_0\right\}, \text{ pour } k \geq 1.$$

Pour $k \geq 1$, sachant que la suite Y_t est décroissante (voir lemme 5.2.6), nous avons

$$\begin{aligned}
\mathbb{E}(Y_{T_k} \mid Y_{T_{k-1}} \geq \rho n) &\geq \mathbb{E}(Y_{T_k} 1_{\{Y_{T_k} \geq \rho n\}} \mid Y_{T_{k-1}} \geq \rho n) \\
&= \mathbb{E}(Y_{T_k} \mid Y_{T_k} \geq \rho n, Y_{T_{k-1}} \geq \rho n)\mathbb{P}\{Y_{T_k} \geq \rho n \mid Y_{T_{k-1}} \geq \rho n\} \\
&= \mathbb{E}(Y_{T_k} \mid Y_{T_k} \geq \rho n)\mathbb{P}\{Y_{T_k} \geq \rho n \mid Y_{T_{k-1}} \geq \rho n\},
\end{aligned}$$

ce qui peut être écrit comme

$$\mathbb{E}(Y_{T_k} \mid Y_{T_k} \geq \rho n) \leq \frac{\mathbb{E}(Y_{T_k} \mid Y_{T_{k-1}} \geq \rho n)}{\mathbb{P}\{Y_{T_k} \geq \rho n \mid Y_{T_{k-1}} \geq \rho n\}}.$$

En utilisant (B.4), nous pouvons écrire

$$\mathbb{E}(Y_{T_k} \mid Y_{T_k} \geq \rho n) \leq \frac{(\mu+4)n}{4\mu\mathbb{P}\{Y_{T_k} \geq \rho n \mid Y_{T_{k-1}} \geq \rho n\}}.$$

La suite Y_t est décroissante (voir lemme 5.2.6) et comme $Y_{T_0} = Y_0 = \|C_0 - L\|^2 \leq K^2$, nous avons, pour $k \geq 0$,

$$\mathbb{E}(Y_{T_k} \mid Y_{T_k} \geq \rho n) \leq \mathbb{E}(Y_{T_0} \mid Y_{T_k} \geq \rho n) \leq K^2.$$

En mettant ensemble les deux inégalités précédentes, nous obtenons pour $k \geq 1$,

$$\begin{aligned}\mathbb{E}(Y_{T_k} \mid Y_{T_k} \geq \rho n) &\leq \min\left\{\frac{(\mu+4)n}{4\mu\mathbb{P}\{Y_{T_k} \geq \rho n \mid Y_{T_{k-1}} \geq \rho n\}}, K^2\right\} \\ &\leq \frac{(\mu+4)n}{4\mu}\min\left\{\frac{1}{\mathbb{P}\{Y_{T_k} \geq \rho n \mid Y_{T_{k-1}} \geq \rho n\}}, \frac{4\mu K^2}{(\mu+4)n}\right\} \\ &= \frac{(\mu+4)n}{4\mu\alpha_k}.\end{aligned}$$

Par définition de α_0, nous avons, pour tout $k \geq 0$,

$$\mathbb{E}(Y_{T_k} \mid Y_{T_k} \geq \rho n) \leq \frac{(\mu+4)n}{4\mu\alpha_k}.$$

En incluant cette inégalité dans la définition de la suite (T_k) donnée par (B.3), pour $k \geq 0$, nous obtenons

$$T_{k+1} \leq T_k + \left\lceil (n-1)\ln\frac{\mu+4}{4\alpha_k}\right\rceil \leq T_k + (n-1)\ln\frac{\mu+4}{4\alpha_k} + 1.$$

En additionnant les différences $T_{i+1} - T_i$, de $i = 0$ à $k-1$, pour $k \geq 1$, cela donne

$$T_k \leq (n-1)\left(k\ln\frac{\mu+4}{4} - \ln\left(\prod_{i=0}^{k-1}\alpha_i\right)\right) + k. \tag{B.6}$$

Pour $k \geq 1$, comme Y_{T_k} est décroissante (voir lemme 5.2.6), nous avons $\mathbb{P}\{Y_{T_k} \geq \rho n \mid Y_{T_{k-1}} < \rho n\} = 0$ et par définition $\alpha_k \leq \mathbb{P}\{Y_{T_k} \geq \rho n \mid Y_{T_{k-1}} \geq \rho n\}$, donc

$$\mathbb{P}\{Y_{T_k} \geq \rho n\} = \mathbb{P}\{Y_{T_k} \geq \rho n \mid Y_{T_{k-1}} \geq \rho n\}\mathbb{P}\{Y_{T_{k-1}} \geq \rho n\} \leq \alpha_k\mathbb{P}\{Y_{T_{k-1}} \geq \rho n\},$$

cela mène à

$$\mathbb{P}\{Y_{T_k} \geq \rho n\} \leq \prod_{i=1}^{k}\alpha_i. \tag{B.7}$$

Comme $\mathbb{P}\{Y_{T_k} \geq \rho n \mid Y_{T_{k-1}} \geq \rho n\} \leq (\mu+4)/(4\rho\mu)$ et $K \geq \sqrt{\rho n}$, nous obtenons par définition $\alpha_k \leq (\mu+4)/(4\rho\mu)$ pour tout $k \geq 0$. Maintenant, pour tout $\delta \in]0;1[$ il existe $k \geq 1$ tel que

$$\prod_{i=1}^{k}\alpha_i < \delta \leq \prod_{i=1}^{k-1}\alpha_i.$$

Nous avons donc, comme $\alpha_0 = (\mu+4)n/(4\mu K^2)$,

$$-\ln\left(\prod_{i=0}^{k-1}\alpha_i\right) = -\ln\left(\prod_{i=1}^{k-1}\alpha_i\right) - \ln(\alpha_0) \leq -\ln(\delta) - \ln(n) - \ln\left(\frac{\mu+4}{4\mu}\right) + 2\ln(K).$$

De plus, comme $\alpha_i \leq (\mu+4)/(4\rho\mu)$, nous avons $\delta \leq \left(\frac{\mu+4}{4\rho\mu}\right)^{k-1}$ ce qui donne

$$k - 1 \leq \frac{-\ln(\delta)}{\ln\left[(\mu+4)/(4\rho\mu)\right]}.$$

En incluant ces résultats dans (B.6) et en utilisant la définition de θ cela donne

$$\begin{aligned}
T_k &\leq (n-1)\left[\left(\frac{\ln\delta}{\ln\left(\frac{\mu+4}{4\rho\mu}\right)}+1\right)\ln\left(\frac{\mu+4}{4}\right)-\ln\delta-\ln\alpha_0\right]+k\\
&= (n-1)\left[\frac{\ln\rho\mu}{\ln\left(\frac{\mu+4}{4\rho\mu}\right)}\ln\delta+\ln\frac{\mu+4}{4}-\ln\frac{\mu+4}{4}+\ln\mu+2\ln K-\ln n\right]+k\\
&= (n-1)\left[-\frac{\ln\rho\mu}{\ln(4\rho\mu)-\ln(\mu+4)}\ln\delta+\ln\mu+2\ln K-\ln n\right]+k\\
&= (n-1)\theta+k\\
&\leq n\theta+\frac{\ln\rho\mu-1}{\ln(4\rho\mu)-\ln(\mu+4)}\ln\delta-\ln\mu-2\ln K+\ln n+1\\
&\leq n\theta.
\end{aligned}$$

Cette dernière inégalité est valide car $\mu > 4(e-1) \Longrightarrow \mu+4 > 4e$ et donc en utilisant (B.5) nous avons $\rho\mu > e$, donc $\ln\delta$ est négatif et son coefficient est positif. Et aussi, par définition de K, nous avons $\ln n - 2\ln K < -\ln\rho$, donc, toujours du fait que $\rho\mu > e$, on a bien $-\ln\mu - 2\ln K + \ln n + 1 \leq 0$.

Nous obtenons finalement, pour $t \geq n\theta$, à partir de (B.7) et en utilisant le fait que Y_t est décroissante

$$\mathbb{P}\{Y_t \geq \rho n\} \leq \mathbb{P}\{Y_{n\theta} \geq \rho n\} \leq \mathbb{P}\{Y_{T_k} \geq \rho n\} \leq \prod_{i=1}^{k}\alpha_i \leq \delta,$$

c'est-à-dire

$$\mathbb{P}\{Y_t < \rho n\} \geq 1-\delta,$$

ce qu'il fallait démontrer. ■

Le théorème suivant nous donne une formule explicite du temps pour que, en partant d'une configuration où le vecteur C_t est dans la boule de rayon $\sqrt{\rho n}$ et de centre L, nous parvenons à une configuration où l'écart maximal entre deux composantes de C_t soit 2, ceci avec une probabilité supérieure à $1-\delta$, sachant que $\delta \in]0;1[$ et $\rho \in]1/4;5/4[$. Ce théorème introduit la donnée λ dont la valeur absolue est la distance de la moyenne ℓ des composantes du vecteur C_t à l'entier le plus proche, plus précisément $\lambda = \ell - \lceil \ell - 0.5 \rceil$.

Théorème 5.2.10 Soit $\delta \in]0;1[$ et $\rho \in]1/4, 5/4[$, si $\|C_0 - L\| \leq \sqrt{\rho n}$, $\ell - \lfloor \ell \rfloor \neq 1/2$ et $\lambda = \ell - \lceil \ell - 0.5 \rceil$ alors nous avons, pour tout $t \geq \tau$,

$$\mathbb{P}\{\|C_t - L\|_\infty \geq 3/2\} \leq \delta,$$

où

$$\tau = \frac{25(n-1)}{12(2-\rho-\lambda^2-|\lambda|)}\left(\ln n - \ln\delta - 2\ln 2 + \ln(\rho+\lambda^2)\right).$$

Preuve du théorème 5.2.10

λ peut aussi s'écrire

$$\lambda = \begin{cases} \ell - \lfloor \ell \rfloor & \text{si } \ell - \lfloor \ell \rfloor < 1/2 \\ \\ \ell - \lceil \ell \rceil & \text{si } \ell - \lfloor \ell \rfloor > 1/2. \end{cases}$$

Nous pouvons noter que λ est positif dans le premier cas et négatif dans le second. Dans chacun des cas nous avons $|\lambda| < 1/2$ et $\ell - \lambda$ est le plus proche entier de ℓ.

Si $\|C_0 - L\| \leq \sqrt{\rho n}$ alors, comme $\|C_t - L\|$ est décroissante, nous avons aussi $\|C_t - L\| \leq \sqrt{\rho n}$, pour tout $t \geq 0$. Il s'ensuit que

$$\|C_t - L\|_\infty \leq \|C_t - L\| \leq \sqrt{\rho n}.$$

Comme $|\lambda| \leq 1/2$, cela veut dire que, pour tout $i \in [\![1, n]\!]$, nous avons

$$-\frac{1}{2} - \sqrt{\rho n} \leq \lambda - \sqrt{\rho n} \leq C_t^{(i)} - \ell + \lambda \leq \lambda + \sqrt{\rho n} \leq \frac{1}{2} + \sqrt{\rho n}.$$

Soit $B = \lceil 1/2 + \sqrt{\rho n} \rceil$. D'après l'inégalité précédente, pour tout $i \in [\![1, n]\!]$, on a $C_t^{(i)} - \ell + \lambda \in [\![-B, B]\!]$. Pour $k \in [\![-B, B]\!]$, nous appelons $\alpha_{k,t}$ le nombre d'agents ayant la valeur $\ell - \lambda + k$ à l'instant t, c'est-à-dire

$$\alpha_{k,t} = \left| \left\{ i \in [\![1, n]\!] \mid C_t^{(i)} - \ell + \lambda = k \right\} \right|. \tag{B.8}$$

Nous pouvons facilement vérifier que

$$\sum_{k=-B}^{B} \alpha_{k,t} = n. \tag{B.9}$$

En plus nous avons, de par la définition de $\alpha_{k,t}$,

$$\sum_{k=-B}^{B} (\ell - \lambda + k)\alpha_{k,t} = \sum_{i=1}^{n} C_t^{(i)} = n\ell,$$

ce qui donne en utilisant (B.9)

$$\sum_{k=-B}^{B} k\alpha_{k,t} = n\lambda. \tag{B.10}$$

De même, encore par définition de $\alpha_{k,t}$, nous avons

$$\sum_{k=-B}^{B} (\ell - \lambda + k)^2 \alpha_{k,t} = \sum_{i=1}^{n} \left(C_t^{(i)} \right)^2 = \|C_t\|^2.$$

Sachant que $\|C_t - L\|^2 = \|C_t\|^2 - n\ell^2$, en utilisant (B.9) et (B.10), nous obtenons

$$\sum_{k=-B}^{B} k^2 \alpha_{k,t} = \|C_t - L\|^2 + n\lambda^2. \tag{B.11}$$

Comme $\|C_t - L\|^2$ est décroissante, en utilisant l'hypothèse $\|C_0 - L\|^2 \leq \rho n$, nous obtenons $\|C_t - L\|^2 \leq \rho n$ et par conséquent

$$\sum_{k=-B}^{B} k^2 \alpha_{k,t} \leq \rho n + n\lambda^2. \tag{B.12}$$

Soit $x \in [\![1, n]\!]$ défini par

$$x = \sum_{k=0}^{B} \alpha_{k,t}.$$

Nous avons alors

$$\sum_{k=-B}^{-1} \alpha_{k,t} = n - x \tag{B.13}$$

et

$$\sum_{k=-B}^{-1} k\alpha_{k,t} \leq \sum_{k=-B}^{-1} -\alpha_{k,t} = -(n - x).$$

En utilisant (B.10), nous obtenons

$$\sum_{k=1}^{B} k\alpha_{k,t} = n\lambda - \sum_{k=-B}^{-1} k\alpha_{k,t} \geq n\lambda + n - x. \tag{B.14}$$

Nous avons aussi en utilisant (B.13) et (B.14)

$$\sum_{k=-B}^{B} k^2 \alpha_{k,t} = \sum_{k=1}^{B} k^2 \alpha_{k,t} + \sum_{k=-B}^{-1} k^2 \alpha_{k,t} \geq \sum_{k=1}^{B} k\alpha_{k,t} - \sum_{k=-B}^{-1} k\alpha_{k,t} \geq 2(n - x) + n\lambda.$$

En combinant cette inégalité avec (B.12) nous obtenons

$$2(n - x) + n\lambda \leq \sum_{k=-B}^{B} k^2 \alpha_{k,t} \leq \rho n + n\lambda^2.$$

Nous pouvons en déduire que

$$x = \sum_{k=0}^{B} \alpha_{k,t} > \frac{(2 - \rho - \lambda^2 + \lambda)n}{2}.$$

En utilisant le même raisonnement pour la somme $\sum_{k=-B}^{0} \alpha_{k,t}$ nous avons

$$\sum_{k=0}^{B} \alpha_{k,t} > \frac{(2 - \rho - \lambda^2 + \lambda)n}{2} \quad \text{et} \quad \sum_{k=-B}^{0} \alpha_{k,t} > \frac{(2 - \rho - \lambda^2 - \lambda)n}{2}. \tag{B.15}$$

Nous pouvons en déduire

$$\sum_{k=0}^{B} \alpha_{k,t} > \frac{(2 - \rho - \lambda^2 - |\lambda|)n}{2} \quad \text{et} \quad \sum_{k=-B}^{0} \alpha_{k,t} > \frac{(2 - \rho - \lambda^2 - |\lambda|)n}{2}. \tag{B.16}$$

Nous introduisons maintenant la suite $(\Phi_t)_{t\geq 0}$ définie par

$$\Phi_t = \sum_{k=2}^{B} k^2 \alpha_{k,t} + \sum_{k=-B}^{-2} k^2 \alpha_{k,t}.$$

D'après (B.12), Φ_t est majorée de cette manière

$$\Phi_t \leq \sum_{k=-B}^{B} k^2 \alpha_{k,t} \leq \rho n + \lambda^2 n. \tag{B.17}$$

Nous introduisons aussi les ensembles H_t^+ et H_t^- définis par

$$H_t^+ = \{i \in [\![1,n]\!] \mid C_t^{(i)} - \ell + \lambda \geq 2\},$$

$$H_t^- = \{i \in [\![1,n]\!] \mid C_t^{(i)} - \ell + \lambda \leq -2\}$$

et nous définissons $H_t = H_t^+ \cup H_t^-$.
En utilisant la définition de α_k (B.8), on vérifie facilement que

$$\Phi_t = \sum_{i \in H_t} \left(C_t^{(i)} - \ell + \lambda\right)^2. \tag{B.18}$$

Soient I_t^+ et I_t^- les ensembles définis par

$$I_t^+ = \{i \in [\![1,n]\!] \mid C_t^{(i)} - \ell + \lambda \geq 0\},$$

$$I_t^- = \{i \in [\![1,n]\!] \mid C_t^{(i)} - \ell + \lambda \leq 0\}.$$

Les relations (B.16) peuvent être réécrites de cette manière

$$|I_t^+| \geq \frac{(2-\rho-\lambda^2-|\lambda|)n}{2} \quad \text{et} \quad |I_t^-| \geq \frac{(2-\rho-\lambda^2-|\lambda|)n}{2}. \tag{B.19}$$

Rappelons que la variable aléatoire X_t, qui donne le couple d'agents qui interagissent à l'instant t, est uniformément distribuée, c'est-à-dire que, pour tout $i,j \in [\![1,n]\!]$, nous avons

$$\mathbb{P}\{X_t = (i,j)\} = \frac{1_{i\neq j}}{n(n-1)}.$$

X_t est la variable aléatoire qui permet de passer de C_t à C_{t+1}, X_t est indépendante de C_t, par contre toutes les autre variables indicées par t comme entre autres I_t^-, Φ_t, $\alpha_{i,t}$, sont entièrement déterminées par C_t.

La principale façon de faire décroître Φ_t est qu'un agent de H_t^+ interagisse avec un agent de I_t^- ou qu'un agent de H_t^- interagisse avec un agent de I_t^+, à l'instant t. Aussi nous allons considérer les événements où un agent de H_t^+ interagit avec un agent de I_t^- et où un agent de H_t^- interagit avec un agent de I_t^+, à l'instant t. Ensuite nous minorerons la décroissance $\Phi_t - \Phi_{t+1}$ dans ces cas, cela permettra d'obtenir une minoration de $\mathbb{E}(\Phi_t - \Phi_{t+1} \mid C_t)$ en fonction de Φ_t.

Soit $E = (H_t^+ \times I_t^-) \cup (I_t^- \times H_t^+) \cup (H_t^- \times I_t^+) \cup (I_t^+ \times H_t^-)$, l'ensemble des interactions "intéressantes", nous allons montrer que, pour ces interactions, la décroissante

de Φ est d'au moins 12/25 fois le poids des états dans Φ_t des noeuds interagissant. Ou exprimé différemment, sur chacune de ces interactions la décroissance $\Phi_{t+1} - \Phi_t$ est supérieure à 12/25 fois la somme des états au carré des nœuds interagissant, éléments de H_t.

Nous introduisons la notation

$$G_t^+ = I_t^+ \setminus H_t^+ = \{i \in [\![1, n]\!] \mid C_t^{(i)} - \ell + \lambda \in \{0, 1\}\},$$

$$G_t^- = I_t^- \setminus H_t^- = \{i \in [\![1, n]\!] \mid C_t^{(i)} - \ell + \lambda \in \{-1, 0\}\}$$

et $G_t = G_t^+ \cup G_t^-$.

Nous allons considérer maintenant la différence $\Phi_t - \Phi_{t+1}$ en fonction des différentes interactions ayant lieu à l'instant t.

Supposons que $X_t = (i, j)$ avec $i \neq j$ et que $X_t \in E$. Nous avons les deux cas suivants.

Cas 1) $(i, j) \in \left(H_t^+ \times G_t^-\right) \cup \left(G_t^- \times H_t^+\right) \cup \left(H_t^- \times G_t^+\right) \cup \left(G_t^+ \times H_t^-\right)$. H_t et G_t étant disjoints, pour simplifier les calculs nous notons a l'élément provenant de H_t et b l'élément de provenant G_t c'est-à-dire

$$a = \left(C_t^{(i)} - \ell + \lambda\right) 1_{\{i \in H_t\}} + \left(C_t^{(j)} - \ell + \lambda\right) 1_{\{j \in H_t\}},$$

$$b = \left(C_t^{(i)} - \ell + \lambda\right) 1_{\{i \in G_t\}} + \left(C_t^{(j)} - \ell + \lambda\right) 1_{\{j \in G_t\}}.$$

Nous avons

$$\Phi_t - \Phi_{t+1} = a^2 - \left(\frac{a + b - 1_{\{a+b \text{ impair}\}}}{2}\right)^2 1_{\{i \in H_{t+1}\}} - \left(\frac{a + b + 1_{\{a+b \text{ impair}\}}}{2}\right)^2 1_{\{j \in H_{t+1}\}}, \tag{B.20}$$

ce qui donne

$$\Phi_t - \Phi_{t+1} \geq a^2 - \left(\frac{a + b - 1_{\{a+b \text{ impair}\}}}{2}\right)^2 - \left(\frac{a + b + 1_{\{a+b \text{ impair}\}}}{2}\right)^2.$$

En distinguant successivement les cas où $a+b$ est pair ou impair, nous obtenons

$$\Phi_t - \Phi_{t+1} \geq \frac{a^2}{2} - b\left(a + \frac{b}{2}\right) - \frac{1_{\{a+b \text{ impair}\}}}{2}. \tag{B.21}$$

Nous allons considérer les cas $|b| = 1$ et $b = 0$ séparément.

Si $|b| = 1$, alors $|a| \geq 2$ et le signe de a est forcément différent du signe de b donc $-ab \geq 2$ donc $\Phi_t - \Phi_{t+1} \geq a^2/2 + 1 \geq 12a^2/25$.

Si $b = 0$, alors nous distinguons 4 cas : a est pair, $|a| = 3$, $|a| = 5$ et $a \geq 5$.

Si a est pair, alors, comme $b = 0$, nous avons à partir de (B.21),

$$\Phi_t - \Phi_{t+1} \geq \frac{a^2}{2} \geq \frac{12a^2}{25}.$$

Si $|a| = 3$, alors nous devons revenir à la relation (B.20)

$$\Phi_t - \Phi_{t+1} = 9 - 4 = 5 \geq \frac{a^2}{2} \geq \frac{12a^2}{25}.$$

Si $|a| = 5$, alors nous avons

$$\Phi_t - \Phi_{t+1} = 25 - 4 - 9 = 12 = \frac{12a^2}{25}.$$

Si a est impair et $|a| \geq 7$, alors nous avons

$$\Phi_t - \Phi_{t+1} = a^2 - \left(\frac{a+b-1}{2}\right)^2 - \left(\frac{a+b+1}{2}\right)^2 = \frac{a^2-1}{2} \geq \frac{12a^2}{25}.$$

Ainsi nous avons montré que si $(i,j) \in \left(H_t^+ \times G_t^-\right) \cup \left(G_t^- \times H_t^+\right) \cup \left(H_t^- \times G_t^+\right) \cup \left(G_t^+ \times H_t^-\right)$, alors

$$\Phi_t - \Phi_{t+1} \geq \frac{12a^2}{25}.$$

Cas 2) $(i,j) \in \left(H_t^+ \times H_t^-\right) \cup \left(H_t^- \times H_t^+\right)$, pour simplifier l'écriture nous allons utliser la notation

$$a = \left(C_t^{(i)} - \ell + \lambda\right) \quad \text{et} \quad b = \left(C_t^{(j)} - \ell + \lambda\right).$$

Nous avons

$$\begin{aligned}\Phi_t - \Phi_{t+1} = a^2 + b^2 &- \left(\frac{a+b-1_{\{a+b \text{ impair}\}}}{2}\right)^2 1_{\{i \in H_{t+1}\}} \\ &- \left(\frac{a+b+1_{\{a+b \text{ impair}\}}}{2}\right)^2 1_{\{j \in H_{t+1}\}}, \end{aligned} \tag{B.22}$$

ce qui donne

$$\begin{aligned}\Phi_t - \Phi_{t+1} &\geq a^2 + b^2 - \left(\frac{a+b-1_{\{a+b \text{ impair}\}}}{2}\right)^2 - \left(\frac{a+b+1_{\{a+b \text{ impair}\}}}{2}\right)^2 \\ &= \frac{a^2}{2} + \frac{b^2}{2} - ab - \frac{1_{\{a+b \text{ impair}\}}}{2}. \end{aligned} \tag{B.23}$$

Par définition de H_t^+ et H_t^-, nous avons $-ab \geq 4$, aussi nous obtenons

$$\Phi_t - \Phi_{t+1} \geq \frac{a^2}{2} + \frac{b^2}{2} \geq \frac{12a^2}{25} + \frac{12b^2}{25}.$$

En mettant ensemble les cas **1)** et **2)**, nous obtenons

$$\begin{aligned}E = &\left(H_t^+ \times G_t^-\right) \cup \left(G_t^- \times H_t^+\right) \cup \left(H_t^- \times G_t^+\right) \cup \left(G_t^+ \times H_t^-\right) \\ &\cup \left(H_t^+ \times H_t^-\right) \cup \left(H_t^- \times H_t^+\right).\end{aligned}$$

Ces six ensembles sont disjoints deux à deux, aussi en utilisant les résultats obtenus dans les cas 1) et 2), et en définissant $\beta_{t,i} = \left(C_t^{(i)} - \ell + \lambda\right)^2$, nous avons

$$\begin{aligned}\sum_{(i,j)\in E} \mathbb{E}(\Phi_t - \Phi_{t+1} \mid C_t, X_t = (i,j)) &\geq \frac{12}{25} \sum_{i\in H_t^+} \sum_{j\in G_t^-} \beta_{t,i} + \frac{12}{25} \sum_{i\in G_t^-} \sum_{j\in H_t^+} \beta_{t,j} \\ &+ \frac{12}{25} \sum_{i\in H_t^-} \sum_{j\in G_t^+} \beta_{t,i} + \frac{12}{25} \sum_{i\in G_t^+} \sum_{j\in H_t^-} \beta_{t,j} + \frac{12}{25} \sum_{i\in H_t^+} \sum_{j\in H_t^-} [\beta_{t,i} + \beta_{t,j}] \\ &+ \frac{12}{25} \sum_{i\in H_t^-} \sum_{j\in H_t^+} [\beta_{t,i} + \beta_{t,j}] \\ = \frac{12}{25} \left[2|G_t^-| \sum_{i\in H_t^+} \beta_{t,i} + 2|G_t^+| \sum_{i\in H_t^-} \beta_{t,i} \right.&+ \left. 2|H_t^+| \sum_{i\in H_t^-} \beta_{t,i} + 2|H_t^-| \sum_{i\in H_t^+} \beta_{t,i} \right].\end{aligned}$$

En observant que $|G_t^-| + |H_t^-| = |I_t^-|$, $|G_t^+| + |H_t^+| = |I_t^+|$ et sachant que $|I_t^-| \geq (2-\rho-\lambda^2-|\lambda|)n/2$ et $|I_t^+| \geq (2-\rho-\lambda^2-|\lambda|)n/2$ (relations (B.19)), nous obtenons

$$\begin{aligned}\sum_{(i,j)\in E} \mathbb{E}(\Phi_t - \Phi_{t+1} \mid C_t, X_t = (i,j)) &\geq \frac{12}{25} \left[2|I_t^-| \sum_{i\in H_t^+} \beta_{t,i} + 2|I_t^+| \sum_{i\in H_t^-} \beta_{t,i} \right] \\ &\geq \frac{12(2-\rho-\lambda^2-|\lambda|)n}{25} \Phi_t. \qquad \text{(B.24)}\end{aligned}$$

Il nous reste à montrer que Φ ne croît pas pour les autres interactions. Plus précisément, nous allons montrer que si, à l'instant t, le couple d'agents interagissant est $(i,j) \notin E$, alors nous avons $\Phi_t - \Phi_{t+1} \geq 0$. La condition $(i,j) \notin E$ à l'instant t peut être décomposée en deux cas disjoints :

- $(i,j) \in \left(H_t^+ \times H_t^+\right) \cup \left(H_t^- \times H_t^-\right)$,
- $\left(C_t^{(i)} - \ell + \lambda = 1 \text{ et } j \in H_t^+\right)$ ou $\left(C_t^{(i)} - \ell + \lambda = -1 \text{ et } j \in H_t^-\right)$ ou $\left(i \in H_t^+ \text{ et } C_t^{(j)} - \ell + \lambda = 1\right)$ ou $\left(i \in H_t^- \text{ et } C_t^{(j)} - \ell + \lambda = -1\right)$.

Si $(i,j) \in \left(H_t^+ \times H_t^+\right)$, alors $(i,j) \in \left(H_{t+1}^+ \times H_{t+1}^+\right)$ et si $(i,j) \in \left(H_t^- \times H_t^-\right)$, alors $(i,j) \in \left(H_{t+1}^- \times H_{t+1}^-\right)$ car, à partir de (5.1), $C_{t+1}^{(i)}$ et $C_{t+1}^{(j)}$ sont les valeurs moyennes de $C_t^{(i)}$ et $C_t^{(j)}$. Donc dans ces cas à partir de la relation (B.18), nous

avons

$$\begin{aligned}\Phi_t - \Phi_{t+1} &= \left(C_t^{(i)} - \ell + \lambda\right)^2 + \left(C_t^{(j)} - \ell + \lambda\right)^2 \\ &\quad - \left(\left\lfloor \frac{C_t^{(i)} + C_t^{(j)}}{2} \right\rfloor - \ell + \lambda\right)^2 - \left(\left\lceil \frac{C_t^{(i)} + C_t^{(j)}}{2} \right\rceil - \ell + \lambda\right)^2 \\ &= \left(C_t^{(i)} - \ell\right)^2 + \left(C_t^{(j)} - \ell\right)^2 \\ &\quad - \left(\left\lfloor \frac{C_t^{(i)} + C_t^{(j)}}{2} \right\rfloor - \ell\right)^2 - \left(\left\lceil \frac{C_t^{(i)} + C_t^{(j)}}{2} \right\rceil - \ell\right)^2 \\ &= \|C_t - L\|^2 - \|C_{t+1} - L\|^2 \geq 0.\end{aligned}$$

Nous considérons maintenant le second cas où $\left(C_t^{(i)} - \ell + \lambda = 1 \text{ et } j \in H_t^+\right)$ ou $\left(C_t^{(i)} - \ell + \lambda = -1 \text{ et } j \in H_t^-\right)$ ou $\left(i \in H_t^+ \text{ et } C_t^{(j)} - \ell + \lambda = 1\right)$ ou $\left(i \in H_t^- \text{ et } C_t^{(j)} - \ell + \lambda = -1\right)$. Pour simplifier la notation, nous définissons

$$a = |C_t^{(i)} - \ell + \lambda| 1_{\{i \in H_t\}} + |C_t^{(j)} - \ell + \lambda| 1_{\{j \in H_t\}}.$$

Nous allons maintenant distinguer trois sous-cas.

- Si $a = 2$, nous avons $\Phi_t - \Phi_{t+1} = 0$ car, à partir de (5.1) l'interaction entre les valeurs 1 et 2 donne 1 et 2, et l'interaction entre -2 et -1 donne -2 et -1.
- Si $a \geq 3$ et a est impair, alors $\Phi_t - \Phi_{t+1} = a^2 - 2\left((a+1)/2\right)^2 = (a^2 - 2a - 1)/2 \geq 0$ pour $a \geq 3 > 1 + \sqrt{2}$.
- Si $a \geq 4$ et a est pair, alors $\Phi_t - \Phi_{t+1} = a^2 - \left((a+2)/2\right)^2 - \left(a/2\right)^2 = (a^2 - 2a - 2)/2 \geq 0$ pour $a \geq 4 \geq 1 + \sqrt{3}$.

Nous pouvons donc écrire

$$\sum_{(i,j) \notin E, i \neq j} \mathbb{E}(\Phi_t - \Phi_{t+1} \mid C_t, X_t = (i,j)) \geq 0. \tag{B.25}$$

Tous les événements $\{X_t = (i,j)\}$ ont la même probabilité, avec $\mathbb{P}\{X_t = (i,j)\} = p_{i,j} = 1_{\{i \neq j\}}/(n(n-1))$, aussi en utilisant les inégalités (B.24) et (B.25), nous obtenons

$$\begin{aligned}\mathbb{E}(\Phi_t - \Phi_{t+1} \mid C_t) &= \sum_{i=1}^{n} \sum_{j=1}^{n} p_{i,j} \mathbb{E}(\Phi_t - \Phi_{t+1} \mid C_t, X_t = (i,j)) \\ &= \frac{1}{n(n-1)} \left(\sum_{(i,j) \in E} \mathbb{E}(\Phi_t - \Phi_{t+1} \mid C_t, X_t = (i,j)) \right. \\ &\qquad \left. + \sum_{(i,j) \notin E, i \neq j} \mathbb{E}(\Phi_t - \Phi_{t+1} \mid C_t, X_t = (i,j)) \right) \\ &\geq \frac{12(2 - \rho - \lambda^2 - |\lambda|)}{25(n-1)} \Phi_t.\end{aligned}$$

Comme $\mathbb{E}(\Phi_t \mid C_t) = \Phi_t$, cela mène à

$$\mathbb{E}(\Phi_{t+1}|C_t) \leq \left(1 - \frac{12(2-\rho-\lambda^2-|\lambda|)}{25(n-1)}\right)\Phi_t$$

et par conséquent

$$\mathbb{E}(\Phi_t) \leq \left(1 - \frac{12(2-\rho-\lambda^2-|\lambda|)}{25(n-1)}\right)^t \Phi_0.$$

Soit τ défini par

$$\tau = \frac{25(n-1)}{12(2-\rho-\lambda^2-|\lambda|)}\left(\ln n - \ln\delta - 2\ln 2 + \ln(\rho+\lambda^2)\right).$$

En utilisant le fait que $\ln(1-x) \leq -x$, pour tout $x \in [0,1[$, nous obtenons, pour tout $t \geq \tau$,

$$\begin{aligned}\left(1 - \frac{12(2-\rho-\lambda^2-|\lambda|)}{25(n-1)}\right)^t &= e^{t\ln(1-12(2-\rho-\lambda^2-|\lambda|)/(25(n-1)))} \\ &\leq e^{-12(2-\rho-\lambda^2-|\lambda|)t/(25(n-1))} \\ &\leq e^{-12(2-\rho-\lambda^2-|\lambda|)\tau/(25(n-1))} \\ &= \frac{4\delta}{(\rho+\lambda^2)n}.\end{aligned}$$

En utilisant l'inégalité de Markov et la relation (B.17) qui donne $\mathbb{E}(\Phi_0) \leq \rho n + \lambda^2$, nous obtenons

$$\mathbb{P}\{\Phi_t \geq 4\} \leq \frac{\mathbb{E}(\Phi_t)}{4} \leq \left(\frac{4\delta}{(\rho+\lambda^2)n}\right)\left(\frac{(\rho+\lambda^2)n}{4}\right) = \delta.$$

Par définition de Φ_t, nous avons $\Phi_t \neq 0 \iff \Phi_t \geq 4$. De plus, en utilisant $|\lambda| < 1/2$, nous avons

$$\begin{aligned}\Phi_t = 0 &\implies \alpha_{t,k} = 0, \text{ pour tout } k \in H_t \\ &\implies -1 \leq C_t^{(i)} - \ell + \lambda \leq 1, \text{ pour tout } i \in [\![1,n]\!] \\ &\implies -1-\lambda \leq C_t^{(i)} - \ell \leq 1-\lambda, \text{ pour tout } i \in [\![1,n]\!] \\ &\implies -3/2 < C_t^{(i)} - \ell < 3/2, \text{ pour tout } i \in [\![1,n]\!] \\ &\implies \|C_t^{(i)} - \ell\|_\infty < 3/2, \text{ pour tout } i \in [\![1,n]\!].\end{aligned}$$

Cela mène à

$$\|C_t^{(i)} - L\|_\infty \geq 3/2 \implies \Phi_t \neq 0 \iff \Phi_t \geq 4,$$

c'est-à-dire

$$\mathbb{P}\{\|C_t^{(i)} - \ell\|_\infty \geq 3/2\} \leq \mathbb{P}\{\Phi_t \neq 0\} = \mathbb{P}\{\Phi_t \geq 4\} \leq \delta,$$

ce qu'il fallait démontrer. ■

Nous présentons maintenant un théorème qui fait la synthèse des théorèmes 5.2.9 et 5.2.11. Ce théorème optimise aussi les paramètres pour donner un résultat numérique du temps nécessaire pour que le protocole basé sur la moyenne avec des entiers arrive à un état où la différence maximale entre deux états soit deux, ceci avec une probabilité supérieure à $1-\delta$.

Théorème 5.2.12 Pour tout $\delta \in]0;1[$, s'il existe une constante K telle que $\|C_0 - L\| \leq K$, alors, pour tout $t \geq n\,(2\ln K + 2.12\ln n - 6.59\ln\delta + 1.88)$, nous avons

$$\mathbb{P}\{\|C_t - L\|_\infty \geq 3/2\} \leq \delta.$$

Preuve du théorème 5.2.12

Nous considérons d'abord le cas où $\ell - \lfloor \ell \rfloor = 1/2$. Comme $\|C_0 - L\|_\infty \leq \|C_0 - L\| \leq K$ et comme

$$(n-1)(2\ln K + \ln n - \ln\delta) \leq n\,(2\ln K + 2.12\ln n - 6.59\ln\delta + 1.88)\,,$$

le théorème 5.2.8 donne

$$\mathbb{P}\{\|C_t - L\|_\infty \neq 1/2\} \leq \delta,$$

pour $t \geq n\,(2\ln K + 2.12\ln n - 6.59\ln\delta + 1.88)$.

Maintenant comme les composantes de C_t sont des entiers et comme $\ell - \lfloor \ell \rfloor = 1/2$, nous avons

$$\mathbb{P}\{\|C_t - L\|_\infty \geq 3/2\} = \mathbb{P}\{\|C_t - L\|_\infty \neq 1/2\} \leq \delta.$$

Considérons maintenant le cas où $\ell - \lfloor \ell \rfloor \neq 1/2$. Soient $\gamma \in]0;1[$ et $\rho \in]1/4;5/4[$. Nous appliquons successivement le théorème 5.2.9 et le théorème 5.2.11 en remplaçant δ par $\gamma\delta$ puis par $(1-\gamma)\delta$. Nous calculons ensuite les valeurs optimales pour les constantes ρ, μ et γ. Nous introduisons la notation

$$\theta_1 = 2\ln K - \ln n + \ln\mu - \frac{\ln\rho\mu}{\ln 4\rho\mu - \ln(\mu+4)}\ln(\gamma\delta)$$

Si $\|C_0 - L\| < \sqrt{\rho n}$ alors nous avons $\|C_0 - L\|^2 < \rho n$ et comme $(\|C_t - L\|^2)_t$ est décroissante (voir lemme 5.2.6), nous obtenons, pour tout $t \geq 0$,

$$\mathbb{P}\{\|C_t - L\|^2 < \rho n\} \geq \mathbb{P}\{\|C_0 - L\|^2 < \rho n\} = 1 \geq 1 - \gamma\delta.$$

Si $\|C_0 - L\| \geq \sqrt{\rho n}$ alors, à partir du théorème 5.2.9, nous obtenons, pour tout $t \geq n\theta_1$, $\mathbb{P}\{\|C_t - L\|^2 \geq \rho n\} \leq \gamma\delta$, ou de manière équivalente

$$\mathbb{P}\{\|C_t - L\|^2 < \rho n\} \geq 1 - \gamma\delta. \tag{B.26}$$

Nous introduisons l'instant τ défini par

$$\tau = n\theta_1 + \frac{25(n-1)}{3(5-4\rho)}\left(\ln n - \ln((1-\gamma)\delta) - 4\ln 2 + \ln(4\rho+1)\right).$$

Nous avons, pour tout $t \geq \tau$,

$$\begin{aligned}\mathbb{P}\{\|C_t - L\|_\infty < 3/2\} &\geq \mathbb{P}\{\|C_t - L\|_\infty < 3/2, \|C_{n\theta_1} - L\|^2 < \rho n\} \\ &= \mathbb{P}\{\|C_t - L\|_\infty < 3/2 \mid \|C_{n\theta_1} - L\|^2 < \rho n\}\mathbb{P}\{\|C_{n\theta_1} - L\|^2 < \rho n\}. \end{aligned} \tag{B.27}$$

Nous avons vu que $\mathbb{P}\{\|C_{n\theta_1} - L\|^2 < \rho n\} \geq 1 - \gamma\delta$. En utilisant le fait que la chaîne de Markov $\{C_t\}$ est homogène et en appliquant le théorème 5.2.11, nous obtenons

$$\begin{aligned}\mathbb{P}\{\|C_t - L\|_\infty < 3/2 \mid \|C_{n\theta_1} - L\|^2 < \rho n\} & \\ &= \mathbb{P}\{\|C_{t-n\theta_1} - L\|_\infty < 3/2 \mid \|C_0 - L\|^2 < \rho n\} \\ &= \mathbb{P}\{\|C_{t-n\theta_1} - L\|_\infty < 3/2 \mid \|C_0 - L\| < \sqrt{\rho n}\} \\ &\geq 1 - (1-\gamma)\delta. \qquad \text{(B.28)}\end{aligned}$$

En incorporant les relations (B.26) et (B.28) à la relation (B.26) nous obtenons pour tout $t \geq \tau$,

$$\mathbb{P}\{\|C_t - L\|_\infty < 3/2\} \geq (1-\gamma\delta)(1-(1-\gamma)\delta) \geq 1 - \delta$$

ou de manière équivalente

$$\mathbb{P}\{\|C_t - L\|_\infty \geq 3/2\} \leq \delta.$$

Le reste de la preuve consite à simplifier l'expression de τ. Nous avons

$$\begin{aligned}\theta_1 &= 2\ln K - \ln n + \ln\mu - \frac{\ln\rho\mu}{\ln 4\rho\mu - \ln(\mu+4)}\ln(\gamma\delta) \\ &= 2\ln K - \ln n + \ln\mu - \frac{\ln\rho\mu}{\ln 4\rho\mu - \ln(\mu+4)}\ln\gamma - \frac{\ln\rho\mu}{\ln 4\rho\mu - \ln(\mu+4)}\ln\delta\end{aligned}$$

et

$$\begin{aligned}\tau &= n\theta_1 + \frac{25(n-1)}{3(5-4\rho)}\left(\ln n - \ln\delta - \ln(1-\gamma) - 4\ln 2 + \ln(4\rho+1)\right) \\ &\leq n\left[2\ln K + \left(\frac{25}{3(5-4\rho)} - 1\right)\ln n - \left(\frac{\ln\rho\mu}{\ln 4\rho\mu - \ln(\mu+4)} + \frac{25}{3(5-4\rho)}\right)\ln\delta\right. \\ &\quad \left. + \ln\mu - \frac{\ln\gamma\ln\rho\mu}{\ln 4\rho\mu - \ln(\mu+4)} - \left(\frac{25}{3(5-4\rho)}\right)\left(\ln(1-\gamma) + 4\ln 2 - \ln(4\rho+1)\right)\right].\end{aligned}$$

En prenant $\rho = 0.5814$, $\mu = 9.2780$ et $\gamma = 0.5270$, nous obtenons

$$\tau = n\left(2\ln K + 2.12\ln n - 6.59\ln\delta + 1.88\right),$$

ce qu'il fallait démontrer. ■

Annexe C

Horloge globale

Cette section de l'annexe est dédiée aux preuves des lemmes du chapitre 6. Ces preuves ont besoin des deux lemmes techniques qui suivent.

Lemme 6.3.1 Pour tout $x \in \mathbb{R}$, nous avons $1 + x \leq e^x$. Pour tout $x \in]-\infty, c]$, nous avons $e^x \leq 1 + x + x^2$, où c est l'unique solution non nulle de l'équation $e^c - 1 - c - c^2 = 0$. La valeur de c vérifie $1.79 < c < 1.8$.

Preuve du lemme 6.3.1. Soit f, une fonction sur $\mathbb{R}$ telle que $f(x) = e^x - x - 1$. En dérivant f, nous obtenons $f'(x) = e^x - 1$. f' est positive sur $[0, +\infty[$ et négative sur $]-\infty, 0]$, donc 0 est le point minimum de f. $f(0) = 0$ donc $\forall x \in \mathbb{R}, f(x) \geq 0$, on a bien $1 + x \leq e^x$ pour tout $x \in \mathbb{R}$.

Soit g, une fonction sur $\mathbb{R}$ telle que $g(x) = e^x - x^2 - x - 1$. En dérivant g, nous obtenons $g'(x) = e^x - 2x - 1$, puis, en dérivant une seconde fois nous obtenons $g''(x) = e^x - 2$. g'' est strictement croissante et s'annule en $\ln 2$. g' est donc décroissante sur $]-\infty, \ln 2]$ et croissante sur $\ln 2, +\infty[$. Les racines de g' sont 0 et une valeur que nous appelons d, on vérifie aisément que $d \in [1.2, 1.3]$. g' est positive sur $]-\infty, 0] \cup [d, +\infty[$ et négative sur $[0, d]$. 0 est un maximum local de g, or $g(0) = 0$ donc g est négative sur $]-\infty, d]$. d est un minimum local de g donc g est négative sur $]-\infty, c]$, c étant la racine de g non nulle. On a donc bien pour $x \in]-\infty, c]$, $e^x \leq 1 + x + x^2$. On vérifie facilement que $c \in]1.79, 1.8[$.

■

Lemme 6.3.2 Soient $u = (u_k)_{k\geq 1}$ et $v = (v_k)_{k\geq 1}$, deux suites monotones de nombres réels et soit $(m_n)_{n\geq 1}$, la suite des valeurs moyennes de la suite v définie, pour $n \geq 1$, par

$$m_n = \frac{1}{n} \sum_{k=1}^{n} v_k.$$

Si les deux suites u et v sont toutes les deux croissantes ou toutes les deux décroissantes, nous avons

$$\sum_{k=1}^{n} u_k v_k \geq m_n \sum_{k=1}^{n} u_k.$$

Si une des deux suites est croissante et l'autre est décroissante alors nous avons

$$\sum_{k=1}^{n} u_k v_k \leq m_n \sum_{k=1}^{n} u_k.$$

Preuve du lemme 6.3.2. Supposons que les deux suites soient croissantes. Soit h l'indice tel que $v_h \leq m_n \leq v_{h+1}$. Comme $\sum_{k=1}^{n}(v_k - m_n) = 0$, nous avons

$$-\sum_{k=1}^{h}(v_k - m_n) = \sum_{k=h+1}^{n}(v_k - m_n) \geq 0$$

et par conséquent

$$\begin{aligned}
\sum_{k=1}^{n} u_k v_k - m_n \sum_{k=1}^{n} u_k &= \sum_{k=1}^{n} u_k (v_k - m_n) \\
&= \sum_{k=1}^{h} u_k (v_k - m_n) + \sum_{k=h+1}^{n} u_k (v_k - m_n) \\
&\geq u_h \sum_{k=1}^{h} (v_k - m_n) + u_{h+1} \sum_{k=h+1}^{n} (v_k - m_n) \\
&= (u_{h+1} - u_h) \sum_{k=h+1}^{n} (v_k - m_n) \\
&\geq 0.
\end{aligned}$$

En multipliant les deux suites par -1, nous obtenons le cas où les deux suites sont décroissantes. Le dernier cas est obtenu en multipliant une seule des suites par -1. ■

Le lemme technique suivant est nécessaire pour prouver les lemmes 6.3.9 et 6.3.10.

Lemme C.0.1 *Soient $a, b, c, d > 0$.*
Si $ad - bc \geq 0$, alors $\dfrac{a+b}{c+d} \leq \dfrac{a}{c}$.
Si $ad - bc \leq 0$ et $c - d > 0$, alors $\dfrac{a-b}{c-d} \leq \dfrac{a}{c}$.

Preuve. Si $ad - bc \geq 0$, alors

$$\frac{a+b}{c+d} - \frac{a}{c} = \frac{ac + bc - ac - ad}{(c+d)c} = \frac{bc - ad}{(c+d)c} \leq 0.$$

Si $ad - bc \leq 0$ et $c - d > 0$, alors

$$\frac{a-b}{c-d} - \frac{a}{c} = \frac{ac - bc - ac + ad}{(c-d)c} = \frac{ad - bc}{(c-d)c} \leq 0,$$

ce qui termine la preuve. ■

Nous supposons, dans le lemme suivant, que l'hypothèse principale du lemme 6.3.7 n'est pas vérifiée, c'est-à-dire nous supposons que $x_{(1-\mu)n} > 0$. Afin de simplifier l'écriture, nous introduisons la notation suivante pour tout $\lambda \in]0,1[$ tel que $\lambda n \in \mathbb{N}$,

$$\Phi_{\leq \lambda n}(t) = \sum_{i=1}^{\lambda n} e^{\alpha x_i(t)}, \qquad \Phi_{>\lambda n}(t) = \sum_{i=\lambda n+1}^{n} e^{\alpha x_i(t)},$$
$$\Psi_{\leq \lambda n}(t) = \sum_{i=1}^{\lambda n} e^{-\alpha x_i(t)} \text{ et } \Psi_{>\lambda n}(t) = \sum_{i=\lambda n+1}^{n} e^{-\alpha x_i(t)}.$$

Lemme 6.3.9 Soient $\alpha, \mu \in]0; 1/2[$ avec $\mu n \in \mathbb{N}$ et $\mu \in]\alpha/(1+\alpha), (1-2\alpha)/(1-\alpha)[$, soit $\mu' \in]0,1[$ avec $\mu' n \in \mathbb{N}$ et $\mu' \in]\mu/(1-\mu), 1/(1+\alpha)[$ et soit $\gamma_1 \in]0,1[$.

Si $x_{(1-\mu)n} > 0$ et $\mathbb{E}(\Delta\Phi(t) \mid x(t)) \geq -(1-\mu'(\alpha+1))\dfrac{\alpha\gamma_1}{n}\Phi(t)$ et $\Phi(t) \geq \lambda_1 \Psi(t)$, alors nous avons

$$\Gamma(t) \leq c_1 n,$$

où

$$c_1 = \left(1 + \frac{1}{\lambda_1}\right) C_1 \left(\frac{C_1}{\mu\lambda_1}\right)^{\mu/((1-\mu)\mu'-\mu)}, \; C_1 = \frac{(1-\mu')(2+\alpha)}{(1-\gamma_1)(1-\mu'(1+\alpha))},$$

$$\text{et } \lambda_1 = \frac{1-\mu-\alpha(2-\mu)}{2\alpha}.$$

La condition $\mu < (1-2\alpha)/(1-\alpha)$ est nécessaire pour que $\lambda_1 > 0$. La valeur de λ_1 sera utilisée dans le théorème 6.3.11. La condition $\mu' > \mu/(1-\mu)$ est nécessaire pour que la puissance utilisée dans le calcul de c_1 soit positive.

Preuve du lemme 6.3.9. Comme $-(\alpha/n - \alpha^2/n^2) \leq 0$, nous avons, en utilisant le lemme 6.3.3

$$\mathbb{E}\left(\Phi(t+1) - \Phi(t) \mid x(t)\right) \leq \left(\alpha + \alpha^2\left(1 - \frac{2}{n}\right)\right) \sum_{i=1}^{n} p_i e^{\alpha x_i(t)} - \left(\frac{\alpha}{n} - \frac{\alpha^2}{n^2}\right)\Phi(t)$$
$$\leq \left(\alpha + \alpha^2\left(1 - \frac{2}{n}\right)\right) \sum_{i=1}^{\mu' n} p_i e^{\alpha x_i(t)}$$
$$+ \left(\alpha + \alpha^2\left(1 - \frac{2}{n}\right)\right) \sum_{i=\mu' n+1}^{n} p_i e^{\alpha x_i(t)} - \left(\frac{\alpha}{n} - \frac{\alpha^2}{n^2}\right)\Phi_{\leq \mu' n}(t).$$

La suite $\left(e^{\alpha x_i(t)}\right)_i$ est décroissante et la suite $(p_i)_i$ est croissante, aussi en prenant successivement

$$m_n = \frac{1}{\mu' n}\sum_{i=1}^{\mu' n} p_i = \frac{\mu' n - 1}{n(n-1)} = \frac{\mu'}{n} - \frac{1-\mu'}{n(n-1)} \leq \frac{1}{n}\left(\mu' - \frac{1-\mu'}{n}\right) \text{ puis}$$
$$m_n = \frac{1}{(1-\mu')n}\sum_{i=\mu' n+1}^{n} p_i = \frac{1}{n}\left(1 + \mu' + \frac{\mu'}{n-1}\right)$$

et en appliquant le lemme 6.3.2 nous obtenons

$$\begin{aligned}
&\mathbb{E}\left(\Delta\Phi(t) \mid x(t)\right) \\
&\quad\leq \left[\left(\mu' - \frac{1-\mu'}{n}\right)\frac{1}{n}\left(\alpha + \alpha^2\left(1-\frac{2}{n}\right)\right) - \left(\frac{\alpha}{n} - \frac{\alpha^2}{n^2}\right)\right]\Phi_{\leq \mu' n}(t) \\
&\qquad + \left(\alpha + \alpha^2\left(1-\frac{2}{n}\right)\right)\frac{1}{n}\left(1+\mu' + \frac{\mu'}{n-1}\right)\Phi_{>\mu' n}(t) \\
&\quad= \left[\left(\mu' - \frac{1-\mu'}{n}\right)\left(1+\alpha\left(1-\frac{2}{n}\right)\right) - \left(1-\frac{\alpha}{n}\right)\right]\frac{\alpha}{n}\Phi_{\leq \mu' n}(t) \\
&\qquad + \left(1+\alpha\left(1-\frac{2}{n}\right)\right)\left(1+\mu' + \frac{\mu'}{n-1}\right)\frac{\alpha}{n}\Phi_{>\mu' n}(t) \\
&\quad= -\left[1-\mu'(\alpha+1) + \frac{1-\mu'}{n} + \alpha\mu'\left(1+\frac{2}{n}\right)\right]\frac{\alpha}{n}\Phi_{\leq \mu' n}(t) \\
&\qquad + \left[1+\alpha - \frac{2\alpha(1+\mu')}{n} + \mu'(1+\alpha) + \frac{\mu'(1+\alpha)}{n-1} - \frac{2\alpha\mu'}{n(n-1)}\right]\frac{\alpha}{n}\Phi_{>\mu' n}(t) \\
&\quad= -\left[1-\mu'(\alpha+1) + \frac{1-\mu'}{n}\right]\frac{\alpha}{n}\Phi_{\leq \mu' n}(t) \\
&\qquad + \left[1+\alpha - \frac{2\alpha(1+\mu')}{n} + \mu'(1+\alpha) + \frac{\mu'(1+\alpha)}{n-1} - \frac{2\alpha\mu'}{n(n-1)}\right]\frac{\alpha}{n}\Phi_{>\mu' n}(t).
\end{aligned}$$

En utilisant le fait que $\Phi_{\leq \mu' n}(t) = \Phi(t) - \Phi_{>\mu' n}(t)$, nous obtenons

$$\begin{aligned}
\mathbb{E}\left(\Phi(t+1) - \Phi(t) \mid x(t)\right) &= -\left[1-\mu'(\alpha+1) + \frac{1-\mu'}{n}\right]\frac{\alpha}{n}\Phi(t) \\
&\quad + \left[2+\alpha + \frac{1-\alpha(2+\mu')}{n-1} - \frac{1-\mu'-2\alpha}{n(n-1)}\right]\frac{\alpha}{n}\Phi_{>\mu' n}(t).
\end{aligned}$$

En utilisant maintenant la deuxième hypothèse qui vérifie

$$\begin{aligned}
\mathbb{E}\left(\Phi(t+1) - \Phi(t) \mid x(t)\right) &\geq -\left(1-\mu'(\alpha+1)\right)\frac{\alpha\gamma_1}{n}\Phi(t) \\
&\geq -\left(1-\mu'(\alpha+1) + \frac{1-\mu'}{n}\right)\frac{\alpha\gamma_1}{n}\Phi(t),
\end{aligned}$$

nous obtenons

$$\begin{aligned}
&\left[2+\alpha + \frac{1-\alpha(2+\mu')}{n-1} - \frac{1-\mu'-2\alpha}{n(n-1)}\right]\frac{\alpha}{n}\Phi_{>\mu' n}(t) \\
&\qquad \geq \left[1-\mu'(\alpha+1) + \frac{1-\mu'}{n}\right]\frac{\alpha(1-\gamma_1)}{n}\Phi(t).
\end{aligned}$$

Notons que la condition $\mu' < 1/(1+\alpha)$ implique que $1-\mu'(\alpha+1) > 0$.

Nous introduisons la notation $B(t) = \sum_{i=1}^{n} \max(0, x_i(t))$. La suite $(x_i(t))_i$ étant décroissante, nous avons, pour tout $\ell, \ell' \in \mathbb{N}$ avec $1 \leq \ell' < \ell \leq n$

$$\sum_{i=\ell'}^{\ell} x_i(t) \leq \frac{(\ell - \ell' + 1)B}{\ell}. \tag{C.1}$$

À partir de (C.1), en prenant $\ell = \ell' = \mu' n$, nous avons

$$x_{\mu' n}(t) \leq B(t)/(\mu' n)$$

En utilisant la décroissance de la suite $(x_i(t))_i$, il s'ensuit

$$\Phi_{>\mu' n}(t) = \sum_{i=\mu' n+1}^{n} e^{\alpha x_i(t)} \leq (1-\mu') n e^{\alpha x_{\mu' n}(t)} \leq (1-\mu') n e^{\alpha B(t)/(\mu' n)}.$$

Cela mène à

$$\begin{aligned}\Phi(t) &\leq \frac{2+\alpha+\dfrac{1-\alpha(2+\mu')}{n-1}-\dfrac{1-\mu'-2\alpha}{n(n-1)}}{(1-\gamma_1)\left(1-\mu'(\alpha+1)+\dfrac{1-\mu'}{n}\right)}\Phi_{>\mu' n} \\ &\leq \frac{(1-\mu')\left(2+\alpha+\dfrac{1-\alpha(2+\mu')}{n-1}-\dfrac{1-\mu'-2\alpha}{n(n-1)}\right)}{(1-\gamma_1)\left(1-\mu'(\alpha+1)-\dfrac{1-\mu'}{n}\right)} n e^{\alpha B(t)/(\mu' n)}.\end{aligned}$$

Nous allons maintenant utiliser le lemme C.0.1. Nous définissons d'abord

$$a = 2+\alpha,\ b = \frac{1-\alpha(2+\mu')}{n-1} - \frac{1-\mu'-2\alpha}{n(n-1)},\ c = 1-\mu'(\alpha+1) \text{ et } d = \frac{1-\mu'}{n}.$$

Nous avons $a, c, d > 0$. Si $b \leq 0$, alors nous avons de manière évidente $\dfrac{a+b}{c+d} \leq \dfrac{a}{c}$. Si $b > 0$, nous obtenons, après quelques calculs algébriques,

$$\begin{aligned}(ad-bc)n(n-1) = &\left[(n-1)(1+3\alpha)-\mu'\right](1-\mu') \\ &+ (n-1)\alpha\mu'(2-\mu'-2\alpha-\alpha\mu') + \mu'\alpha(1-\alpha\mu').\end{aligned}$$

Comme $\alpha < 1/2$, la condition $\mu' < 1/(1+\alpha)$ implique que $\mu' < 2(1-\alpha)/(1+\alpha)$ qui à son tour implique que $2-\mu'-2\alpha-\alpha\mu' > 0$. Nous en déduisons que $ad - bc > 0$ et en utilisant le lemme C.0.1, nous obtenons $\dfrac{a+b}{c+d} \leq \dfrac{a}{c}$. Cela mène à

$$\Phi(t) \leq \frac{n(1-\mu')(2+\alpha)}{(1-\gamma_1)(1-\mu'(\alpha+1))} e^{\alpha B(t)/(\mu' n)}.$$

Nous introduisons la notation $C_1 = \dfrac{(1-\mu')(2+\alpha)}{(1-\gamma_1)(1-\mu'(\alpha+1))}$. Nous pouvons écrire

$$\Phi(t) \leq C_1 n e^{\alpha B(t)/(\mu' n)}.$$

La fonction exponentielle étant convexe, l'inégalité de Jensen donne

$$\Psi(t) \geq \Psi_{>(1-\mu)n}(t) = \sum_{i=(1-\mu)n+1}^{n} e^{-\alpha x_i(t)} \geq \mu n \exp\left(-\frac{\alpha \sum_{i=(1-\mu)n+1}^{n} x_i(t)}{\mu n}\right).$$

Considérons la somme $\sum_{i=(1-\mu)n+1}^{n} x_i(t)$ et rappelons que la suite $(x_i(t))_i$ est décroissante. Comme $x_{(1-\mu)n}(t) > 0$, cette somme contient toutes les valeurs négatives de $x_i(t)$ dont la somme est égale à $-B(t)$. Soit r le nombre de $x_i(t)$ positifs de la somme $\sum_{i=(1-\mu)n+1}^{n} x_i(t)$. Notons que $r \in [\![0, \mu n - 1]\!]$. Nous avons, en utilisant C.1,

$$\sum_{i=(1-\mu)n+1}^{n} x_i(t) = -B(t) + \sum_{i=(1-\mu)n+1}^{(1-\mu)n+r} x_i(t) \leq -B(t) + \frac{rB(t)}{(1-\mu)n+r}.$$

La fonction f définie, pour $r \in [0, \mu n]$, par $f(r) = rB(t)/((1-\mu)n + r)$ étant croissante, nous avons

$$\frac{rB(t)}{(1-\mu)n+r} = f(r) \leq f(\mu n) = \mu B(t),$$

ce qui mène à

$$\sum_{i=(1-\mu)n+1}^{n} x_i(t) \leq -B(t) + \mu B(t) = -(1-\mu)B(t)$$

et ainsi, nous obtenons

$$\Psi(t) \geq \mu n \exp\left(-\frac{\alpha \sum_{i=(1-\mu)n+1}^{n} x_i(t)}{\mu n}\right) \geq \mu n e^{\alpha(1-\mu)B(t)/(\mu n)}.$$

L'hypothèse $\Phi(t) \geq \lambda_1 \Psi(t)$ donne

$$C_1 n e^{\alpha B/(\mu' n)} \geq \Phi(t) \geq \lambda_1 \Psi(t) \geq \lambda_1 \mu n e^{\alpha(1-\mu)B(t)/(\mu n)},$$

ce qui à son tour donne

$$\frac{C_1}{\lambda_1 \mu} \geq \exp\left(\frac{\alpha B(t)}{n}\left(\frac{1-\mu}{\mu} - \frac{1}{\mu'}\right)\right) = \exp\left(\frac{\alpha B(t)}{\mu' n}\left(\frac{(1-\mu)\mu' - \mu}{\mu}\right)\right),$$

c'est-à-dire

$$e^{\alpha B(t)/(\mu' n)} \leq \left(\frac{C_1}{\lambda_1 \mu}\right)^{\frac{\mu}{(1-\mu)\mu' - \mu}}.$$

Nous arrivons finalement à

$$\begin{aligned}\Gamma(t) = \Phi(t) + \Psi(t) &\leq \left(1 + \frac{1}{\lambda_1}\right)\Phi(t) \leq \left(1 + \frac{1}{\lambda_1}\right) C_1 n e^{\alpha B/(\mu' n)} \\ &\leq \left(1 + \frac{1}{\lambda_1}\right) C_1 n \left(\frac{C_1}{\lambda_1 \mu}\right)^{\frac{\mu}{(1-\mu)\mu' - \mu}},\end{aligned}$$

ce qu'il fallait démontrer. ■

Nous supposons dans le lemme suivant que l'hypothèse principale du lemme 6.3.8 n'est pas vérifiée, c'est-à-dire que nous supposons que $x_{\rho n} < 0$.

Lemme 6.3.10 Soient $\alpha, \rho \in]0;1/2[$ avec $\rho n \in \mathbb{N}$ et $\rho \in]\alpha/(1-\alpha); 1/(1+\alpha)[$, soit $\rho' \in]\rho/(1-\rho), (1-2\alpha)/(1-\alpha)[$ avec $\rho' n \in \mathbb{N}$ et soit $\gamma_2 \in]0,1[$.

Si $x_{\rho n} < 0$ et $\mathbb{E}(\Delta\Psi(t) \mid x(t)) \geq -[1-2\alpha-\rho'(1-\alpha)]\,\dfrac{\alpha\gamma_2}{n}\Psi(t)$ et $\Psi(t) \geq \lambda_2\Phi(t)$, alors nous avons

$$\Gamma(t) \leq c_2 n,$$

où

$$c_2 = \left(1+\frac{1}{\lambda_2}\right) C_2 \left(\frac{C_2}{\rho\lambda_2}\right)^{\rho/((1-\rho)\rho'-\rho)}, \; C_2 = \frac{(1-\rho')\,(2-2\alpha-\rho'(1-\alpha))}{(1-\gamma_2)\,(1-2\alpha-\rho'(1-\alpha))},$$

$$\text{et } \lambda_2 = \frac{1-\rho(1+\alpha)}{2\alpha}.$$

La condition $\rho < 1/(1+\alpha)$ est nécessaire pour que $\lambda_2 > 0$. La valeur de λ_2 sera utilisée dans le théorème 6.3.11. La condition $\rho' > \rho/(1-\rho)$ est nécessaire pour que la puissance utilisée dans le calcul de c_2 soit positive.

Preuve du lemme 6.3.10. En utilisant le lemme 6.3.5 et comme $-\alpha+\alpha^2(1-2/n) \leq 0$, nous avons

$$\begin{aligned}
\mathbb{E}\left(\Psi(t+1)-\Psi(t) \mid x(t)\right) &\leq \left(-\alpha+\alpha^2\left(1-\frac{2}{n}\right)\right)\sum_{i=1}^{n} p_i e^{-\alpha x_i(t)} + \left(\frac{\alpha}{n}+\frac{\alpha^2}{n^2}\right)\Psi(t) \\
&\leq \left(-\alpha+\alpha^2\left(1-\frac{2}{n}\right)\right)\sum_{i=(1-\rho')n}^{n} p_i e^{-\alpha x_i(t)} \\
&\quad + \left(1+\frac{\alpha}{n}\right)\frac{\alpha}{n}\Psi_{>(1-\rho')n}(t) + \left(1+\frac{\alpha}{n}\right)\frac{\alpha}{n}\Psi_{\leq(1-\rho')n}(t).
\end{aligned}$$

La suite $\left(\left(-\alpha+\alpha^2(1-2/n)\right)e^{\alpha x_i(t)}\right)_i$ est décroissante et la suite$(p_i)_i$ est croissante, aussi en prenant m_n tel que

$$m_n = \frac{1}{\rho' n}\sum_{i=(1-\rho')n+1}^{n} p_i = \frac{1}{n}\left(2-\rho'+\frac{1-\rho'}{n-1}\right) \geq \frac{2-\rho'}{n}$$

et en appliquant le lemme 6.3.2 nous obtenons

$$\begin{aligned}
\mathbb{E}\left(\Psi(t+1)-\Psi(t) \mid x(t)\right) &\leq m_n\left(-\alpha+\alpha^2\left(1-\frac{2}{n}\right)\right)\Psi_{>(1-\rho')n}(t) \\
&\quad + \left(1+\frac{\alpha}{n}\right)\frac{\alpha}{n}\Psi_{>(1-\rho')n}(t) + \left(1+\frac{\alpha}{n}\right)\frac{\alpha}{n}\Psi_{\leq(1-\rho')n}(t) \\
&\leq \left[(2-\rho')\left(-1+\alpha\left(1-\frac{2}{n}\right)\right)+1+\frac{\alpha}{n}\right]\frac{\alpha}{n}\Psi_{>(1-\rho')n}(t) \\
&\quad + \left(1+\frac{\alpha}{n}\right)\frac{\alpha}{n}\Psi_{\leq(1-\rho')n}(t) \\
&= -\left[1-2\alpha-\rho'(1-\alpha)+\frac{\alpha(3-2\rho')}{n}\right]\frac{\alpha}{n}\Psi_{>(1-\rho')n}(t) \\
&\quad + \left(1+\frac{\alpha}{n}\right)\frac{\alpha}{n}\Psi_{\leq(1-\rho')n}(t).
\end{aligned}$$

En utilisant le fait $\Psi_{>(1-\rho')n}(t) = \Psi(t) - \Psi_{\leq(1-\rho')n}(t)$, nous obtenons

$$\begin{aligned}\mathbb{E}\left(\Delta\Psi(t) \mid x(t)\right) \leq &- \left[1 - 2\alpha - \rho'(1-\alpha) + \frac{\alpha(3-2\rho')}{n}\right]\frac{\alpha}{n}\Psi(t) \\ &+ \left[2 - 2\alpha - \rho'(1-\alpha) + \frac{\alpha(4-2\rho')}{n}\right]\frac{\alpha}{n}\Psi_{\leq(1-\rho')n}(t).\end{aligned}$$

En utilisant la deuxième hypothèse, nous avons

$$\begin{aligned}\mathbb{E}\left(\Delta\Psi(t) \mid x(t)\right) &\geq -\left[1 - 2\alpha - \rho'(1-\alpha)\right]\frac{\alpha\gamma_2}{n}\Psi(t) \\ &\geq -\left[1 - 2\alpha - \rho'(1-\alpha) + \frac{\alpha(3-2\rho')}{n}\right]\frac{\alpha\gamma_2}{n}\Psi(t)\end{aligned}$$

et par conséquent, nous obtenons

$$\begin{aligned}&\left[2 - 2\alpha - \rho'(1-\alpha) + \frac{\alpha(4-2\rho')}{n}\right]\frac{\alpha}{n}\Psi_{\leq(1-\rho')n}(t) \\ &\qquad\geq (1-\gamma_2)\left[1 - 2\alpha - \rho'(1-\alpha) + \frac{\alpha(3-2\rho')}{n}\right]\frac{\alpha}{n}\Psi(t).\end{aligned}$$

Notons que les conditions $\rho' < (1-2\alpha)/(1-\alpha)$ et $\alpha < 1/2$ impliquent que $1 - 2\alpha - \rho'(1-\alpha) > 0$.

Comme pour la démonstration du théorème 6.3.9, nous introduisons la notation $B(t) = \sum_{i=1}^{n} \max(0, x_i(t))$. La suite $(x_i(t))_i$ étant décroissante, nous avons, pour tout $\ell, \ell' \in \mathbb{N}$ avec $1 \leq \ell < \ell' \leq n$

$$-\frac{(\ell'-\ell)B}{n-\ell} \leq \sum_{i=\ell+1}^{\ell'} x_i(t). \tag{C.2}$$

À partir de (C.2), en prenant $\ell = \ell' = (1-\rho')n$, nous avons

$$x_{(1-\rho')n}(t) \geq -B(t)/(\rho' n)$$

En utilisant la décroissance de la suite $(x_i(t))_i$, il s'ensuit

$$\Psi_{\leq(1-\rho')n}(t) = \sum_{i=1}^{(1-\rho')n} e^{-\alpha x_i(t)} \leq (1-\rho')n e^{-\alpha x_{(1-\rho')n}(t)} \leq (1-\rho')n e^{\alpha B(t)/(\rho' n)}.$$

Ce qui mène à

$$\begin{aligned}\Psi(t) &\leq \frac{\left[2 - 2\alpha - \rho'(1-\alpha) + \frac{\alpha(4-2\rho')}{n}\right]}{(1-\gamma_2)\left[1 - 2\alpha - \rho'(1-\alpha) + \frac{\alpha(3-2\rho')}{n}\right]}\Psi_{\leq(1-\rho')n} \\ &\leq \frac{(1-\rho')\left[2 - 2\alpha - \rho'(1-\alpha) + \frac{\alpha(4-2\rho')}{n}\right]}{(1-\gamma_2)\left[1 - 2\alpha - \rho'(1-\alpha) + \frac{\alpha(3-2\rho')}{n}\right]} n e^{\alpha B/(\rho' n)}.\end{aligned}$$

Nous pouvons maintenant utiliser le lemme C.0.1. Soient $a = 1$, $b = \alpha/n$, $c = 1 - 2\alpha - \rho'(1-\alpha)$ et $d = \alpha(3-2\rho')/n$. Nous avons $a, b, c, d > 0$ et $ad - bc = \alpha(2-\rho')(1+\alpha)/n \geq 0$, aussi en utilisant le lemme C.0.1, nous obtenons $\dfrac{a+b}{c+d} \leq \dfrac{a}{c}$, c'est-à-dire

$$\begin{aligned}\frac{2-2\alpha-\rho'(1-\alpha)+\dfrac{\alpha(4-2\rho')}{n}}{1-2\alpha-\rho'(1-\alpha)+\dfrac{\alpha(3-2\rho')}{n}} &= 1+\frac{1+\dfrac{\alpha}{n}}{1-2\alpha-\rho'(1-\alpha)+\dfrac{\alpha(3-2\rho')}{n}}\\ &\leq 1+\frac{1}{1-2\alpha-\rho'(1-\alpha)}\\ &= \frac{2-2\alpha-\rho'(1-\alpha)}{1-2\alpha-\rho'(1-\alpha)}.\end{aligned}$$

Cela mène à

$$\Psi(t) \leq \frac{(1-\rho')\,(2-2\alpha-\rho'(1-\alpha))}{(1-\gamma_2)\,(1-2\alpha-\rho'(1-\alpha))} n e^{\alpha B(t)/(\mu' n)}.$$

En introduisant la notation $C_2 = \dfrac{(1-\rho')\,(2-2\alpha-\rho'(1-\alpha))}{(1-\gamma_2)\,(1-2\alpha-\rho'(1-\alpha))}$, nous pouvons écrire

$$\Phi(t) \leq C_2 n e^{\alpha B(t)/(\rho' n)}.$$

La fonction exponentielle étant convexe, l'inégalité de Jensen donne

$$\Phi(t) \geq \Phi_{\leq \rho n}(t) = \sum_{i=1}^{\rho n} e^{\alpha x_i(t)} \geq \rho n \exp\left(\frac{\alpha \sum_{i=1}^{\rho n} x_i(t)}{\rho n}\right).$$

Considérons la somme $\sum_{i=1}^{\rho n} x_i(t)$ et rappelons que la suite $(x_i(t))_i$ est décroissante. Comme $x_{\rho n}(t) < 0$, la somme contient toutes les valeurs positives du vecteur $x(t)$ dont la somme est égale à $B(t)$, et au moins une valeur négative $x_i(t)$. Soit r le nombre de valeurs négatives de la somme $\sum_{i=1}^{\rho n} x_i(t)$, notons que $r \in [\![1, \rho n - 1]\!]$. Nous avons, en utilisant (C.2),

$$\sum_{i=1}^{\rho n} x_i(t) = B(t) + \sum_{i=\rho n - r}^{\rho n} x_i(t) \geq B(t) - \frac{r B(t)}{(1-\rho)n + r}.$$

La fonction g définie, pour $r \in [1, \rho n]$, par $g(r) = -rB/((1-\rho)n + r)$ étant décroissante, nous avons

$$-\frac{r B(t)}{(1-\rho)n + r} = g(r) \geq g(\rho n) = -\rho B(t),$$

ce qui mène à

$$\sum_{i=1}^{\rho n} x_i(t) \geq B(t) - \rho B(t) = (1-\rho) B(t)$$

et aussi, nous obtenons

$$\Phi(t) \geq \rho n \exp\left(\frac{\alpha \sum_{i=1}^{\rho n} x_i(t)}{\rho n}\right) \geq \rho n e^{\alpha(1-\rho)B(t)/(\rho n)}.$$

L'hypothèse $\Psi(t) \geq \lambda_2 \Phi(t)$ donne

$$C_2 n e^{\alpha B/(\rho' n)} \geq \Psi(t) \geq \lambda_2 \Phi(t) \geq \lambda_2 \rho n e^{\alpha(1-\rho)B(t)/(\rho n)},$$

ce qui à son tour donne

$$\frac{C_2}{\lambda_2 \rho} \geq \exp\left(\frac{\alpha B(t)}{n}\left(\frac{1-\rho}{\rho} - \frac{1}{\rho'}\right)\right) = \exp\left(\frac{\alpha B(t)}{\rho' n}\left(\frac{(1-\rho)\rho' - \rho}{\rho}\right)\right),$$

c'est-à-dire

$$e^{\alpha B(t)/(\rho' n)} \leq \left(\frac{C_2}{\lambda_2 \rho}\right)^{\frac{\rho}{(1-\rho)\rho' - \rho}}.$$

Finalement nous arrivons à

$$\begin{aligned}\Gamma(t) = \Phi(t) + \Psi(t) &\leq \left(1 + \frac{1}{\lambda_2}\right)\Psi(t) \leq \left(1 + \frac{1}{\lambda_2}\right) C_2 n e^{\alpha B/(\rho' n)} \\ &\leq \left(1 + \frac{1}{\lambda_2}\right) C_2 n \left(\frac{C_2}{\lambda_2 \rho}\right)^{\frac{\rho}{(1-\rho)\rho' - \rho}},\end{aligned}$$

ce qu'il fallait démontrer. ■

Théorème 6.3.11 Soient $\alpha, \mu, \rho \in]0, 1/2[$ avec $\mu n, \rho n \in \mathbb{N}$, $\mu \in (\alpha/(1+\alpha), (1-2\alpha)/(1-\alpha))$ et $\rho \in (\alpha/(1-\alpha), 1/(1+\alpha))$. Soit $\mu' \in (\mu/(1-\mu), 1/(1+\alpha))$ avec $\mu' n \in \mathbb{N}$ et soit $\rho' \in (\rho/(1-\rho), (1-2\alpha)/(1-\alpha))$ avec $\rho' n \in \mathbb{N}$. Soient $\gamma_1, \gamma_2 \in]0; 1[$. Alors nous avons

$$\mathbb{E}\left(\Gamma(t+1) \mid x(t)\right) \leq \left(1 - c_4 \frac{\alpha}{n}\right)\Gamma(t) + c_3,$$

où

$$\begin{aligned}c_4 = \min\Big\{&\mu - \alpha(1-\mu), \rho - \alpha(1+\rho), \gamma_1\left(1 - \mu'(\alpha+1)\right), \frac{\alpha\left(1 - \mu - \alpha(2-\mu)\right)}{1 - \mu(1-\alpha)}, \\ &\gamma_2\left(1 - 2\alpha - \rho'(1-\alpha)\right), \frac{\alpha\left(1 - \rho(1+\alpha)\right)}{1 - \rho(1-\alpha) + 2\alpha}\Big\}\end{aligned}$$

et

$$c_3 = \max\left\{\alpha\left(1 + \alpha + \rho(1+\rho)\right), \alpha(1-\mu)(2-\mu), (\alpha + c_4)\alpha c_1, \alpha + \alpha^2, (\alpha + c_4)\alpha c_2\right\},$$

dans lequel

$$c_1 = \left(1 + \frac{1}{\lambda_1}\right) C_1 \left(\frac{C_1}{\mu\lambda_1}\right)^{\mu/((1-\mu)\mu' - \mu)},\ C_1 = \frac{(1-\mu')(2+\alpha)}{(1-\gamma_1)(1-\mu'(1+\alpha))},$$

$$\lambda_1 = \frac{1 - \mu - \alpha(2-\mu)}{2\alpha}$$

et

$$c_2 = \left(1 + \frac{1}{\lambda_2}\right) C_2 \left(\frac{C_2}{\rho\lambda_2}\right)^{\rho/((1-\rho)\rho' - \rho)},\ C_2 = \frac{(1-\rho')(2 - 2\alpha - \rho'(1-\alpha))}{(1-\gamma_2)(1 - 2\alpha - \rho'(1-\alpha))},$$

$$\lambda_2 = \frac{1 - \rho(1+\alpha)}{2\alpha}.$$

Preuve du théorème 6.3.11. La preuve consiste en l'analyse des trois cas suivants.

- **Cas 1 :** $x_{\rho n} \geq 0$ et $x_{(1-\mu)n} \leq 0$
- **Cas 2 :** $x_{(1-\mu)n} > 0$
- **Cas 3 :** $x_{\rho n} < 0$.

Cas 1 : Supposons $x_{\rho n} \geq 0$ et $x_{(1-\mu)n} \leq 0$. Nous pouvons alors utiliser les lemmes 6.3.7 et 6.3.8. En ajoutant les inégalités (6.9) et (6.10), nous obtenons

$$\mathbb{E}\left(\Gamma(t+1) \mid x(t)\right) \leq \left(1 - \frac{a\alpha}{n}\right)\Gamma(t) + b \leq \left(1 - \frac{c_4\alpha}{n}\right)\Gamma(t) + c_3,$$

où

$$a = \min\left(\mu - \alpha(1-\mu), \rho - \alpha(1+\rho)\right) \geq c_4 \text{ et } b = \alpha\left(1 + \alpha + \rho(1+\rho)\right) \leq c_3.$$

Cas 2 : Supposons que $x_{(1-\mu)n} > 0$. Nous devons alors considérer les trois sous-cas suivants.

- **Cas 2.1 :** $\mathbb{E}\left(\Delta\Phi(t) \mid x(t)\right) \geq -\left(1 - \mu'(\alpha+1)\right)\frac{\alpha\gamma_1}{n}\Phi(t)$ et $\Phi(t) \geq \lambda_1\Psi(t)$,
- **Cas 2.2 :** $\mathbb{E}\left(\Phi(t+1) - \Phi(t) \mid x(t)\right) < -\left(1 - \mu'(\alpha+1)\right)\frac{\alpha\gamma_1}{n}\Phi(t)$,
- **Cas 2.3 :** $\Phi(t) < \lambda_1\Psi(t)$.

Cas 2.1 : Nous supposons que

$$\mathbb{E}\left(\Delta\Phi(t) \mid x(t)\right) \geq -\left(1 - \mu'(\alpha+1) - \frac{\alpha}{n}\right)\frac{\alpha\gamma_1}{n}\Phi(t) \text{ et } \Phi(t) \geq \lambda_1\Psi(t).$$

Nous pouvons alors appliquer le lemme 6.3.9, qui donne $\Gamma(t) \leq c_1 n$.

En additionnant les inégalités obtenues dans les corollaires 6.3.4 et 6.3.6, nous obtenons

$$\mathbb{E}\left(\Gamma(t+1) - \Gamma(t) \mid x(t)\right) \leq \frac{\alpha^2}{n}\left(1 - \frac{1}{n}\right)\Gamma(t) \leq \alpha^2 c_1,$$

et donc, en utilisant le fait que $c_1 \geq \Gamma(t)/n$,

$$\begin{aligned}
\mathbb{E}\left(\Gamma(t+1) \mid x(t)\right) &\leq \Gamma(t) + \alpha^2 c_1 \\
&= \Gamma(t) - c_4\alpha c_1 + \alpha^2 c_1 + c_4\alpha c_1 \\
&\leq \left(1 - c_4\frac{\alpha}{n}\right)\Gamma(t) + (\alpha + c_4)\alpha c_1 \\
&\leq \left(1 - c_4\frac{\alpha}{n}\right)\Gamma(t) + c_3.
\end{aligned}$$

Cas 2.2 : Nous supposons que $\mathbb{E}\left(\Delta\Phi(t) \mid x(t)\right) < -\left(1 - \mu'(\alpha+1)\right)\frac{\alpha\gamma_1}{n}\Phi(t)$.

Comme $\mu, \rho \in]0, 1/2[$, nous avons $\rho < 1 - \mu$. La suite $(x_i(t))_i$ étant décroissante, nous avons $x_{\rho n}(t) \geq x_{(1-\mu)n}(t) > 0$. Par conséquent nous pouvons appliquer le lemme 6.3.8. En ajoutant l'inégalité précédente à celle du lemme 6.3.8 nous obtenons

$$\begin{aligned}
\mathbb{E}\left(\Delta\Gamma(t) \mid x(t)\right) &= \mathbb{E}\left(\Delta\Phi(t) \mid x(t)\right) + \mathbb{E}\left(\Delta\Psi(t) \mid x(t)\right) \\
&\leq -\left(1 - \mu'(\alpha+1)\right)\frac{\alpha\gamma_1}{n}\Phi - \left(\rho - \alpha(1+\rho)\right)\frac{\alpha}{n}\Psi(t) + \alpha\rho(1+\rho) \\
&\leq -\min\left\{\gamma_1\left(1 - \mu'(\alpha+1)\right), \left(\rho - \alpha(1+\rho)\right)\right\}\frac{\alpha}{n}\Gamma(t) + \alpha\rho(1+\rho),
\end{aligned}$$

ce qui donne

$$\begin{aligned}
\mathbb{E}\left(\Gamma(t+1) \mid x(t)\right) &\leq \left(1 - \min\left\{\gamma_1\left(1 - \mu'(\alpha+1)\right), \left(\rho - \alpha(1+\rho)\right)\right\}\frac{\alpha}{n}\right)\Gamma(t) \\
&\qquad + \alpha\rho(1+\rho) \\
&\leq \left(1 - \min\left\{\gamma_1\left(1 - \mu'(\alpha+1)\right), \left(\rho - \alpha(1+\rho)\right)\right\}\frac{\alpha}{n}\right)\Gamma(t) \\
&\qquad + \alpha\left(1 + \alpha + \rho(1+\rho)\right) \\
&\leq \left(1 - c_4\frac{\alpha}{n}\right)\Gamma(t) + c_3.
\end{aligned}$$

Cas 2.3 : Nous supposons que $\Phi < \lambda_1\Psi$, avec $\lambda_1 = \dfrac{1 - \mu - \alpha(2-\mu)}{2\alpha}$. Nous utilisons ici le Collaire 6.3.4 et le lemme 6.3.8 dans lequel nous initialisons $\rho = 1 - \mu$. Nous obtenons, après quelques calculs algébriques,

$$\begin{aligned}
&\mathbb{E}\left(\Delta\Gamma(t) \mid x(t)\right) = \mathbb{E}\left(\Delta\Phi(t) \mid x(t)\right) + \mathbb{E}\left(\Delta\Psi(t) \mid x(t)\right) \\
&\leq \frac{\alpha^2}{n}\left(1 - \frac{1}{n}\right)\Phi(t) - \left[\rho - \alpha(1+\rho) + \frac{\alpha(1+2\rho)}{n}\right]\frac{\alpha}{n}\Psi(t) + \alpha\rho(1+\rho) \\
&\leq \frac{\alpha^2}{n}\left(1 - \frac{1}{n}\right)\lambda_1\Psi(t) - \left[1 - \mu - \alpha(2-\mu) + \frac{\alpha(3-2\mu)}{n}\right]\frac{\alpha}{n}\Psi(t) + \alpha(1-\mu)(2-\mu) \\
&= \left[-\frac{1 - \mu - \alpha(2-\mu)}{2} - \frac{1 - \mu + \alpha(4-3\mu)}{2n}\right]\frac{\alpha}{n}\Psi(t) + \alpha(1-\mu)(2-\mu) \\
&\leq -\frac{1 - \mu - \alpha(2-\mu)}{2}\frac{\alpha}{n}\Psi(t) + \alpha(1-\mu)(2-\mu) \\
&= -\lambda_1\frac{\alpha^2}{n}\Psi(t) + \alpha(1-\mu)(2-\mu).
\end{aligned}$$

En notant que $\Phi(t) \leq \lambda_1\Psi(t) \Longrightarrow \Phi(t) + \Psi(t) \leq (1+\lambda_1)\Psi(t) \Longrightarrow \Psi(t) \geq \dfrac{\Gamma(t)}{1+\lambda_1}$, nous obtenons

$$\begin{aligned}
\mathbb{E}\left(\Gamma(t+1) - \Gamma(t) \mid x(t)\right) &\leq -\frac{\lambda_1}{1+\lambda_1}\frac{\alpha^2}{n}\Gamma(t) + \alpha(1-\mu)(2-\mu) \\
&= -\left(\frac{\alpha\left(1 - \mu - \alpha(2-\mu)\right)}{1 - \mu(1-\alpha)}\right)\frac{\alpha}{n}\Gamma(t) + \alpha(1-\mu)(2-\mu),
\end{aligned}$$

c'est-à-dire

$$\begin{aligned}
\mathbb{E}\left(\Gamma(t+1) \mid x(t)\right) &\leq \left(1 - \left(\frac{\alpha\left(1 - \mu - \alpha(2-\mu)\right)}{1 - \mu(1-\alpha)}\right)\frac{\alpha}{n}\right)\Gamma(t) + \alpha(1-\mu)(2-\mu) \\
&\leq \left(1 - c_4\frac{\alpha}{n}\right)\Gamma(t) + c_3.
\end{aligned}$$

Cas 3 : Nous supposons que $x_{\rho n} < 0$. Nous devons alors considérer les trois sous-cas suivants.

- **Cas 3.1 :** $\mathbb{E}\left(\Delta\Psi(t) \mid x(t)\right) \geq -\left(1 - 2\alpha - \rho'(1-\alpha)\right)\frac{\alpha\gamma_2}{n}\Psi(t)$ et $\Psi(t) \geq \lambda_2\Phi(t)$
- **Cas 3.2 :** $\mathbb{E}\left(\Psi(t+1) - \Psi(t) \mid x(t)\right) < -\left(1 - 2\alpha - \rho'(1-\alpha)\right)\frac{\alpha\gamma_2}{n}\Psi(t)$
- **Cas 3.3 :** $\Psi(t) < \lambda_2\Phi(t)$.

Cas 3.1 : Nous supposons que

$$\mathbb{E}\left(\Delta\Psi(t) \mid x(t)\right) \geq -\left(1 - 2\alpha - \rho'(1-\alpha)\right)\frac{\alpha\gamma_2}{n}\Psi(t) \text{ et } \Psi(t) \geq \lambda_2\Phi(t).$$

Nous pouvons alors appliquer le lemme 6.3.10, qui donne $\Gamma(t) \leq c_2 n$.

En additionnant les inégalités obtenues dans les corollaires 6.3.4 et 6.3.6, nous obtenons

$$\mathbb{E}\left(\Gamma(t+1) - \Gamma(t) \mid x(t)\right) \leq \frac{\alpha^2}{n}\left(1 - \frac{1}{n}\right)\Gamma(t) \leq \alpha^2 c_2,$$

et alors, en utilisant le fait que $c_2 \geq \Gamma(t)/n$,

$$\begin{aligned}
\mathbb{E}\left(\Gamma(t+1) \mid x(t)\right) &\leq \Gamma(t) + \alpha^2 c_2 \\
&= \Gamma(t) - c_4\alpha c_2 + \alpha^2 c_2 + c_4\alpha c_2 \\
&\leq \left(1 - c_4\frac{\alpha}{n}\right)\Gamma(t) + (\alpha + c_4)\alpha c_2 \\
&\leq \left(1 - c_4\frac{\alpha}{n}\right)\Gamma(t) + c_3.
\end{aligned}$$

Cas 3.2 :

Nous supposons que $\mathbb{E}\left(\Delta\Psi(t) \mid x(t)\right) < -\left(1 - 2\alpha - \rho'(1-\alpha)\right)\frac{\alpha\gamma_2}{n}\Psi(t)$.

Comme $\mu, \rho \in]0; 1/2[$, nous avons $\rho < 1 - \mu$. La suite $(x_i(t))_i$ étant décroissante, nous avons $x_{(1-\mu)n}(t) \leq x_{\rho n}(t) < 0$. Nous pouvons donc appliquer le lemme 6.3.7. En additionnant l'inégalité précédente et celle du lemme 6.3.7 nous obtenons

$$\begin{aligned}
\mathbb{E}\left(\Delta\Gamma(t) \mid x(t)\right) &= \mathbb{E}\left(\Delta\Phi(t) \mid x(t)\right) + \mathbb{E}\left(\Delta\Psi(t) \mid x(t)\right) \\
&\leq -\left(\mu - \alpha(1-\mu)\right)\frac{\alpha}{n}\Phi(t) - \gamma_2\left(1 - 2\alpha - \rho'(1-\alpha)\right)\frac{\alpha}{n}\Psi(t) + \alpha + \alpha^2 \\
&\leq -\min\left\{\mu - \alpha(1-\mu), \gamma_2\left(1 - 2\alpha - \rho'(1-\alpha)\right)\right\}\frac{\alpha}{n}\Gamma(t) + \alpha + \alpha^2,
\end{aligned}$$

ce qui donne

$$\begin{aligned}
&\mathbb{E}\left(\Gamma(t+1) \mid x(t)\right) \\
&\qquad \leq \left(1 - \min\left\{\mu - \alpha(1-\mu), \gamma_2\left(1 - 2\alpha - \rho'(1-\alpha)\right)\right\}\right)\frac{\alpha}{n}\Gamma(t) + \alpha + \alpha^2 \\
&\qquad \leq \left(1 - c_4\frac{\alpha}{n}\right)\Gamma(t) + c_3.
\end{aligned}$$

Cas 3.3 : Nous supposons que $\Psi(t) < \lambda_2\Phi(t)$, avec $\lambda_2 = \dfrac{1-\rho(1+\alpha)}{2\alpha}$. Nous utilisons ici le corollaire 6.3.4 et le lemme 6.3.7 dans lequel $\mu = 1-\rho$. Nous obtenons, après quelques calculs algébriques,

$$\begin{aligned}
&\mathbb{E}\left(\Delta\Gamma(t) \mid x(t)\right) = \mathbb{E}\left(\Delta\Phi(t) \mid x(t)\right) + \mathbb{E}\left(\Delta\Psi(t) \mid x(t)\right)\\
&\leq -\left(\mu - \alpha(1-\mu) + \frac{\alpha(1-2\mu)}{n}\right)\frac{\alpha}{n}\Phi(t) + \frac{\alpha^2}{n}\left(1-\frac{1}{n}\right)\Psi(t) + \alpha + \alpha^2\left(1-\frac{2}{n}\right)\\
&\leq -\left(1-\rho(1+\alpha) - \frac{\alpha(1-2\rho)}{n}\right)\frac{\alpha}{n}\Phi(t) + \frac{\alpha^2}{n}\left(1-\frac{1}{n}\right)\lambda_2\Phi(t) + \alpha + \alpha^2\\
&= -\frac{1}{2}\left(1-\rho(1+\alpha) + \frac{1-\rho(1-3\alpha)-2\alpha}{n}\right)\frac{\alpha}{n}\Phi(t) + \alpha + \alpha^2.
\end{aligned}$$

Notons que pour $\rho \in]0, 1/2[$, nous avons $(1-\rho)/(3-2\rho) > 1/2$ ce qui implique que $\alpha < (1-\rho)/(3-2\rho)$ ce qui est équivalent à $1-\rho(1-3\alpha)-2\alpha > 0$. Cela donne

$$\mathbb{E}\left(\Gamma(t+1) - \Gamma(t) \mid x(t)\right) \leq -\left(\frac{1-\rho(1+\alpha)}{2}\right)\frac{\alpha}{n}\Phi(t) + \alpha + \alpha^2 = -\frac{\lambda_2\alpha^2}{n}\Phi(t) + \alpha + \alpha^2.$$

En notant que $\Psi(t) \leq \lambda_2\Phi(t) \Longrightarrow \Phi(t) + \Psi(t) \leq (1+\lambda_2)\Phi(t) \Longrightarrow \Phi(t) \geq \dfrac{\Gamma(t)}{1+\lambda_2}$, nous obtenons

$$\begin{aligned}
\mathbb{E}\left(\Gamma(t+1) - \Gamma(t) \mid x(t)\right) &\leq -\frac{\lambda_2\alpha^2}{(1+\lambda_2)n}\Gamma(t) + \alpha + \alpha^2\\
&= -\left(\frac{\alpha(1-\rho(1+\alpha))}{1-\rho(1-\alpha)+2\alpha}\right)\frac{\alpha}{n}\Gamma(t) + \alpha + \alpha^2,
\end{aligned}$$

c'est-à-dire

$$\begin{aligned}
\mathbb{E}\left(\Gamma(t+1) \mid x(t)\right) &\leq \left(1 - \left(\frac{\alpha(1-\rho(1+\alpha))}{1-\rho(1-\alpha)+2\alpha}\right)\frac{\alpha}{n}\right)\Gamma(t) + \alpha + \alpha^2\\
&\leq \left(1 - c_4\frac{\alpha}{n}\right)\Gamma(t) + c_3,
\end{aligned}$$

ce qui termine la preuve. ■

Liste des publications

- Yves Mocquard, Emmanuelle Anceaume, James Aspnes, Yann Busnel, Bruno Sericola, "Counting with population protocols". *14th IEEE International Symposium on Network Computing and Applications (NCA), Boston, MA, U.S.A. 2015.*

- Yves Mocquard,"Compter avec les protocoles de population", *Présentation au 11ème Atelier en Evaluation de Performances Toulouse 2016.*

- Yves Mocquard, Emmanuelle Anceaume, Bruno Sericola "Optimal Proportion Computation with Population Protocols". *15th IEEE International Symposium on Network Computing and Applications (NCA), Cambridge, MA, U.S.A. 2016.*

- Yves Mocquard, Samantha Robert, Bruno Sericola, Emmanuelle Anceaume "Analysis of the Propagation Time of a Rumour in Large-scale Distributed Systems". *15th IEEE International Symposium on Network Computing and Applications (NCA), Cambridge, MA, U.S.A. 2016.* The best student paper award.

- Yves Mocquard, Bruno Sericola, Emmanuelle Anceaume, "Probabilistic Analysis of Counting Protocols in Large-scale Asynchronous and Anonymous Systems ". *16th IEEE International Symposium on Network Computing and Applications (NCA), Cambridge, MA, U.S.A. 2017.*

- Yves Mocquard, Bruno Sericola, Emmanuelle Anceaume, "Balanced allocation and global clock in population protocols : an accurate analysis". *SIROCCO : 25th Colloquium on Structural Information and Communication Complexity, Ma'ale HaHamisha, Israël 2018.*

- Yves Mocquard, Bruno Sericola, Emmanuelle Anceaume, "Probabilistic Analysis of Rumor Spreading Time". *INFORMS Journal On Computing (IJOC), 2018.*

- Yves Mocquard, Bruno Sericola, Emmanuelle Anceaume, "Population Protocols with Convergence Detection". *17th IEEE International Symposium on Network Computing and Applications (NCA), Cambridge, MA, U.S.A. 2018.*

Résumé

Nous nous situons dans le contexte du modèle des protocoles de population. Ce modèle, introduit en 2004 par Angluin et al., fournit les bases théoriques pour analyser les propriétés émergeant d'un système constitué d'agents anonymes interagissant deux à deux. Dans ce cadre, nous analysons en profondeur quatre protocoles : le protocole de diffusion, de moyenne avec des entiers, de moyenne avec des réels et le protocole d'horloge. En ce qui concerne le protocole de diffusion, notre analyse fournit une expression précise et simple de la queue de distribution du temps de diffusion. Nous analysons aussi en profondeur le comportement asymptotique de la distribution quand n tend vers l'infini. En ce qui concerne le protocole de moyenne avec des entiers, nous démontrons que le protocole converge en un temps parallèle de $O(\log n)$ vers un état où la différence maximale entre deux valeurs est égale à 2. Ce résultat nous permet de prouver l'optimalité en espace et en temps de nos protocoles de proportion et de comptage. En ce qui concerne le protocole de moyenne avec des réels, par l'utilisation de la norme 4, nous réduisons considérablement les constantes des bornes de convergence. En ce qui concerne le protocole d'horloge, nous explicitons les constantes des bornes. Ensuite, nous construisons un protocole de proportion avec détection de convergence, qui utilise nos résultats sur la diffusion, la proportion et l'horloge. Nous montrons également que ce protocole de détection de convergence peut s'appliquer à tout protocole dont on connaît explicitement une borne du temps de convergence avec probabilité élevée.

Mots clés : protocoles de population, diffusion de rumeur, boules et urnes, moyenne, probabilités.

Abstract

We are in the context of the population protocols model. This model, introduced in 2004 by Angluin et al., provides the theoretical basis for analyzing the properties emerging from a system consisting of anonymous agents interacting in pairs. In this framework, we analyze in depth four protocols: the spreading protocol, average with integers, average with reals and the clock protocol. Regarding the spreading protocol, our analysis provides a precise and simple expression of the spreading time distribution tail. We also analyze in depth the asymptotic behavior of the distribution when n tends to infinity. Regarding the average protocol with integers, we demonstrate that the protocol converges in a parallel time of $O(\log n)$ to a state where the maximum difference between two values is 2. This result allows us to prove the space and time optimality of our proportion and counting protocols. Regarding the average protocol with reals, using the 4-norm, we significantly reduce the constants of the convergence time bounds. Regarding the clock protocol, we explain the constants of the bounds. Next, we build proportion protocol with convergence detection, which uses our results on diffusion, proportion, and clock. We also show that this convergence detection protocol can be applied to any protocol which is explicitly known as a convergence time bound with high probability.

Keywords : population protocols, rumor spreading, balls-and-bins, average, probabilities.